

The research program of the Center for Economic Studies (CES) produces a wide range of theoretical and empirical economic analyses that serve to improve the statistical programs of the U.S. Bureau of the Census. Many of these analyses take the form of CES research papers. The papers are intended to make the results of CES research available to economists and other interested parties in order to encourage discussion and obtain suggestions for revision before publication. The papers are unofficial and have not undergone the review accorded official Census Bureau publications. The opinions and conclusions expressed in the papers are those of the authors and do not necessarily represent those of the U.S. Bureau of the Census. Republication in whole or part must be cleared with the authors.

INDUSTRY LEARNING ENVIRONMENTS AND THE HETEROGENEITY OF FIRM PERFORMANCE

by

Natarajan Balasubramanian *
UCLA Anderson School of Management

and

Marvin B. Lieberman *
UCLA Anderson School of Management

CES 06-29

December, 2006

All papers are screened to ensure that they do not disclose confidential information. Persons who wish to obtain a copy of the paper, submit comments about the paper, or obtain general information about the series should contact Sang V. Nguyen, Editor, [Discussion Papers](#), Center for Economic Studies, Washington Plaza II, Room 206, Bureau of the Census, Washington, DC 20233-6300, (301-763-1882) or INTERNET address snnguyen@ces.census.gov.

Abstract

This paper characterizes inter-industry heterogeneity in rates of learning-by-doing and examines how industry learning rates are connected with firm performance. Using data from the Census Bureau and Compustat, we measure the industry learning rate as the coefficient on cumulative output in a production function. We find that learning rates vary considerably among industries and are higher in industries with greater R&D, advertising, and capital intensity. More importantly, we find that higher rates of learning are associated with wider dispersion of Tobin's q and profitability among firms in the industry. Together, these findings suggest that learning intensity represents an important characteristic of the industry environment.

Keywords: Learning, Firm Heterogeneity, RBV, Productivity

* The research in this paper was conducted while the authors were Census Bureau research associates at the California Census Research Data Center (CCRDC). Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the Bureau of the Census. The results presented in this paper have been screened to ensure that no confidential data are revealed. We thank Becky Acosta at the CCRDC for her help with the project proposal and prompt processing of our disclosure requests. We have benefitted immensely from insightful comments by Emin Dinlersoz, Susan Cohen, participants at the CRES-Gort and ACAC conferences and three anonymous reviewers for the 2006 AOM annual meeting. Natarajan Balasubramanian thanks Jagadeesh Sivadasan for numerous discussions on this study, Dan Akerberg, Roberto Alvarez, Michael Darby, Hugo Hopenhayn, Raphael Lam, Jeongsik Lee, Gabriel Natividad and Mariko Sakakibara for their constructive comments and Hyowook Chiang, John Haltiwanger and Ron Jarmin for kindly sharing various data. The CCRDC generously provided support to Natarajan Balasubramanian and covered the data access fees.

INTRODUCTION

This paper examines a key aspect of the firm's industry environment – the importance of learning from direct operating experience or “learning-by-doing” in determining business performance. Industries vary considerably in the relevance of learning-by-doing to firm performance. In some industries, products and processes may be relatively simple, or entrepreneurs and managers may be able to leverage external sources (e.g. specialized technology suppliers, consultants or competitors' employees) to acquire knowledge about their business operations. Other industry environments may not support such acquisition of knowledge or may involve complex, knowledge-intensive processes and products, thereby constraining firms to improve performance largely through direct experience. In such environments, learning-by-doing may significantly affect firm performance.

In this study, we focus on the importance of accumulated experience in the production process as a measure of the importance of learning-by-doing in an industry (“industry learning intensity”). Further, we examine how differences in industry learning intensity are associated with business performance. Using plant-level data from the US Census Bureau on over 55,000 manufacturing plants from 1973 to 2000, we estimate the industry learning rate as the coefficient on prior cumulative output in a production function. Applying these industry learning rates to firm data from Compustat, we find that the cross-sectional variation in business performance within an industry, as measured by the inter-percentile range (10th to 90th) of firm q and firm profitability, is much greater in industries with higher learning intensity. Such findings suggest that learning intensity is an important characteristic of the industry environment that has been omitted from prior studies of firm and industry performance, given data limitations.

This study draws from a long line of literature on organizational learning and “learning curves”, which have been studied since the 1930s. The “learning curve” or the empirical

relationship between unit cost of production and operating experience has been estimated for numerous industries – e.g. ships (Rapping, 1965; Thornton and Thompson, 2001), chemicals (Lieberman, 1984) and semiconductors (Gruber, 1994 and Gruber, 2000) – and generally appear to follow a “power-law” i.e. the unit cost of production decreases at a decreasing rate with increasing experience, typically measured as prior cumulative output. While most studies have found that performance improves as organizations accumulate operating experience, the rate of learning appears to vary across industries. In a review of 22 field studies on learning-by-doing, Dutton and Thomas (1984) found that “progress ratios” (i.e. the percentage decrease in unit costs with a doubling of cumulative experience) ranged from 55% to over 100%. Learning rates have been found to vary within an industry and even within subunits of the same firm. Hayes and Clark (1986) examined productivity in factories and found that learning curves varied significantly across factories within the same company. In an analysis of cardiac surgery departments implementing a new technology for minimally invasive cardiac surgery, Pisano, Bohmer and Edmondson (2001) found that the learning curve slope varied significantly across organizations. While these studies have demonstrated that learning rates vary significantly among organizations and industries, these studies have drawn from limited data sets and have not attempted to characterize differences in learning rates across a broad range of industries.

This study also links to another long line of papers, mostly in the Structure-Conduct-Performance literature that examines how industry factors affect firm performance. Many empirical studies have examined how the variables such as industry structure, R&D intensity and advertising intensity affect firm performance (See Schmalensee (1989) for a review.). The industry learning intensity or the role of direct experience has not been studied (empirically) as a variable that could affect firm performance. This is a bit surprising given that a number of studies, both theoretical and empirical, have argued that the learning curve has implications for competitive strategy and may be used to generate “first-mover advantages” (e.g. Spence, 1981;

Lieberman, 1987).

This study makes two contributions to the existing literature on learning. First, it provides a broad-brush characterization of plant-level learning-by-doing in over 100 SIC-3 digit industries in the US manufacturing sector. This characterization reinforces findings in prior studies that industries vary considerably in their learning rates. In addition, it provides reasonably comparable quantitative estimates of the importance of learning from direct experience across industries. Most prior studies have focused on a single product or service, largely due to non-availability of longitudinal data across industries. In this study, we use a large sample drawn from US Census Bureau data that span the entire US manufacturing sector. We adopt a production function approach and measure the industry learning intensity as the coefficient on prior cumulative output in a production function. This approach is approximately equivalent to the traditional unit-cost learning curve and provides a reasonably uniform measure of learning rates across industries, albeit subject to some limitations. We find that the industry learning rate displays considerable heterogeneity across industries and that it is positively correlated with industry capital-labor ratio, R&D intensity and advertising intensity, even after controlling for joint industry-year fixed effects or plant fixed effects. These correlations are consistent with the intuitive notion that learning-by-doing may be more important in industries with greater process complexity, tacit knowledge and differentiated products.

More importantly, our study demonstrates that industry learning intensity, as measured by plant-level learning-by-doing, has robust relationships with firm performance. In particular, we find that the cross-sectional heterogeneity of firm performance within an industry, as measured by the inter-percentile range of firm profits or firm q , is higher in industries with higher rates of learning. In other words, in such industries, the difference between the “best” and the “worst” (conditional on survival) firms is considerably higher. Taken together, our findings add to the existing literature by introducing industry learning intensity as an important component of

the industry environment that may explain competitive heterogeneity.

MEASURING INTENSITY OF LEARNING-BY-DOING

The traditional approach to measuring learning-by-doing for a product is to estimate a power-law function of the following form:

$$C=AX^{-\lambda} \quad (1)$$

Where:

C is the unit cost of the product;

A is a constant;

X is a measure of experience, typically prior cumulative production; and

$\lambda > 0$ is the rate of learning-by-doing.

This formulation is purely empirical and is a reduced form representation of the various processes of learning from direct experience. The disadvantage of this approach is that it requires detailed cost and production quantity data, which are not easily available for a large number of firms. Our method for measuring learning-by-doing follows Bahk and Gort (1993) and is a variant of the traditional approach. Bahk and Gort (1993) incorporate learning-by-doing within a production function and estimate the coefficients using data from individual manufacturing plants. Following their approach, we can write:

$$Y_{ijt} = \Phi_{jt} (K_{ijt})^{\alpha_j} (L_{ijt})^{\beta_j} (X_{ijt})^{\lambda_j} v_{ijt} \quad (2)$$

Where

Y is the current period real value added, measured as *real revenues less real materials expenses*;

Φ is a constant (explained below);

K and L are real capital stock and quantity of labor, respectively;

X is prior cumulative output, a measure of experience;

α , β and λ are all positive and less than 1;

v is a plant-specific term (explained below);

and subscripts i , j and t refer to plant ' i ', industry ' j ' and year ' t ' respectively.

This formulation is an extension of the widely used Cobb-Douglas production function. In addition to the usual inputs of capital and labor, prior operating experience is considered an “input” into the production process in the sense that a higher level of operating experience increases output for any given level of capital and labor. Hence, λ , the coefficient on prior experience denotes the industry learning intensity.

We can interpret the learning coefficients obtained from this approach in two ways. First, the coefficient λ can be interpreted in a straightforward manner as the importance of learning (from direct experience) in the production process. A higher value of λ implies a greater role for accumulated experience in the production process. We could also interpret learning to be an improvement in “productivity” resulting from experience. Productivity (or more precisely, Total Factor Productivity) as defined in the economics literature is a measure of the efficiency of *physical* resource use. Hence, firms with higher productivity have the capability to generate more or better quality output using the same amount of physical resources. There are two physical resources considered above, capital and labor. So, we could define productivity of a plant as $P_{ijt} = \Phi_{jt}(X_{ijt})^\lambda v_{ijt}$, which is simply the term on the right hand side of Equation 2 excluding the physical resource terms. The second term, X_{ijt}^λ , is the increase in productivity resulting from accumulated direct operating experience and reflects learning-by-doing. The coefficient λ here is a measure of the importance of direct experience in productivity improvement. This definition also enables us to isolate learning-by-doing from other sources of productivity improvement. The first term, Φ_{jt} , captures any industry-wide improvements in productivity (subscript ' j ' refers to industry). This may occur, for instance, because of innovations in the equipment used in the industry, or because of improved practices that become available to all firms in the industry.

The last term, v_{ijt} , captures any improvements in productivity resulting from firm-specific factors other than learning-by-doing.

Like the traditional learning curve, this approach is purely empirical and does not delve into the mechanisms of learning or even changes in firm behavior as a result of learning. Rather, it attempts to measure learning by attributing observed changes in firm performance to an observable proxy for prior experience. Though a very simple and stylized representation of the complex learning processes at play, we believe that the coefficient λ so obtained can reasonably be interpreted as the importance of prior experience X in the production process. It also offers a number of other advantages, some of which are specific to our context.

First, our study is set in the US manufacturing sector, and it stands to reason that manufacturing processes would be important in determining overall firm performance. Hence, the notion of a “production function” makes intuitive sense and focusing on the importance of experience in the production process or on productivity improvement as a measure of learning is meaningful. Another advantage of this approach is that it controls for efficiency gains resulting from economies of scale. A traditional learning curve includes only the cumulative output, which could easily proxy for the scale of production (Argote, 1999, p.16). By including current levels of physical inputs in the specification, the production function approach controls for the possibility that economies of scale (which is a relation between *current* output and current inputs) rather than learning-by-doing (which depends on *past* output) is driving improvements. As explained above, this approach also allows us to control for the possibility that improvements in manufacturing processes are resulting from industry-wide improvements in technology rather than due to direct experience. This approach also has the nice property that under some reasonable assumptions, it is approximately equivalent to the traditional unit cost learning curve. Finally, compared to a traditional learning curve formulation, this approach involves variables that are more easily available. The main disadvantage is that these variables are usually available only at the plant level and not for individual products. Hence, the learning

estimates obtained using this approach represent an average learning rate across products manufactured within a plant.¹

Other approaches to measuring learning-by-doing could be adopted. For instance, firm (or plant) age has been used as a measure of experience in prior studies. The main advantage of using age is that it is an easy-to-understand variable and is less affected by issues of serial correlation than cumulative output. However, the big disadvantage is that using age ignores all the heterogeneity within any age cohort (for instance, not all 5-year old firms are alike, even within an industry). Moreover, prior studies (Lieberman, 1984; Argote, 1999, p. 15) have found that cumulative output is a better measure of experience than time (or age). Hence, in this paper, we prefer prior cumulative output to age as the measure of experience.

HETEROGENEITY IN LEARNING-BY-DOING

Learning-by-doing is generally considered to be the result of organizational search for better routines combined with trial and error experimentation (Levitt and March, 1988; Muth, 1986). Within such a framework, the need for learning from experience as well as the resulting rate of learning-by-doing depend on several factors that necessitate and/or alter the success of organizational search and experimentation processes. The complexity of organizational processes involved is perhaps the most intuitive. Other things being equal, more complex tasks will require and elicit higher investments in learning from experience. The argument is similar for complex products and equipment used by the firm. If the products produced by the firm are complex, it is likely that the firm's consumers go through a learning process and suggest improvements, which if incorporated may manifest as learning-by-doing (von Hippel, 1988). A third factor is the nature of knowledge involved. If the knowledge generated from learning can

¹ As we explain later, we select our sample in a way to reduce the possibility of very different products being manufactured in the same plant.

be easily codified, the need for learning from experience is reduced. On the other hand, if most of the knowledge is tacit, it must be generated by “learning on the job”. A closely related factor is the uncertainty of outcomes and the extent of exogenous “noise” in the processes. If the outcome of a certain process is affected by factors beyond the control of and only incompletely observable by agents involved in the process, the agents will need to experiment many times before they can distinguish between productive and unproductive tasks. In addition to complexity, the opportunities to imitate and learn from other firms also affect the need for learning from own experience. For instance, in a market with very differentiated products, the knowledge acquired by one firm may only be incidentally relevant to other firms. In such a market then, firms may have to rely on their own experience to develop their resources and capabilities.

It is difficult to obtain good empirical measures of process complexity, tacitness of knowledge and product differentiation for such a large sample. Hence, we rely on rather simple proxies that are easily available. Industries with greater capital-intensity have been associated with greater process complexity (Lieberman 1984). Another measure is industry wages – highly skilled workers may be required in industries with greater process complexity and knowledge tacitness. Similarly, it is reasonable to believe that R&D-intensive industries have more complex processes and involve a higher degree of knowledge tacitness. One could also reasonably argue that R&D and advertising-intensive industries are likely to be more differentiated than industries with little R&D or advertising. Following the line of reasoning outlined above, we would then expect high learning-by-doing in such industries. To summarize, we expect a positive relationship between the importance of experience in an industry and the industry capital-labor ratio, wages, R&D and advertising intensity.

INDUSTRY LEARNING-RATE AND HETEROGENEITY OF FIRM PERFORMANCE

There are numerous reasons to expect a positive association between the industry learning intensity and the inter-firm heterogeneity of performance within that industry. First, the learning process is uncertain and complex. The probabilistic nature of the underlying search processes means that similar investments by equally capable agents are unlikely to result in the same degree of success. Moreover, the learning process is fraught with numerous behavioral imperfections and mistakes (Levitt and March, 1988; Levinthal and March, 1993) that further increase the uncertainty of outcomes. These uncertainties and mistakes inherent in the learning process naturally serve to make each firm's evolutionary path, and hence, its bundle of resources and capabilities, distinctive. Hence, as learning becomes more important to firm performance, the resource based view (RBV), which argues that competitive heterogeneity in an industry arises from heterogeneity of firm resources and capabilities (Barney, 1991; Peteraf, 1993), would suggest that we should observe more performance variations across firms. Uncertainty in the learning process may also increase heterogeneity of firm performance by exaggerating the ambiguity regarding a rival's resources and capabilities. If firms could completely observe their rival's learning process, they could then imitate them, potentially reducing competitive heterogeneity. However, with uncertain learning, even though some aspects of the learning process may be observable, rivals may not be able to accurately identify the links between cause and effect, leading to "uncertain imitability" (Lippman and Rumelt, 1989).²

² Along similar lines, Rivkin (2000) uses an agent-based simulation model to suggest that in settings with more complex *strategies* (which is likely if learning is important), imitation efforts are less likely to be successful and consequently, that the "distribution of firm performance will be broader".

Another distinctive aspect of the learning process that can potentially increase competitive heterogeneity is its dynamic character. Learning by doing takes time and hence, firms explicitly consider how today's decisions may affect tomorrow's performance. Issues such as path dependence and time compression diseconomies become more important, thus increasing differences between firms (Dierickx and Cool, 1989). Another possibility is that firms are willing to incur short-term losses to achieve longer-term gains from learning. This may be particularly true for younger firms that are still accumulating experience. In an industry with no learning, profits may remain stable over time and firms with losses in the first one or two years of their life will exit. On the other hand, if profits can possibly increase with time due to learning, firms may be able to offset the initial losses with subsequent gains. Hence, there may be more loss making firms in industries with high learning and we may observe increased competitive heterogeneity.

A third aspect of learning by doing that may increase heterogeneity of firm performance is its self-reinforcing character. Suppose firms enter the market with different capabilities. Then, at entry, the "high capability" firms will typically choose a larger scale of operations and will be more efficient than their "low capability" counterparts. Since firm capabilities are generally persistent, "high capability" firms will have higher level of cumulative output than "low capability" firms at any given time (assuming they entered at the same time), which in turn generates an additional performance advantage for "high capability" firms in the presence of learning-by-doing. Obviously, this self-reinforcing mechanism will be stronger in high learning industries and hence, firm performance will tend to be more heterogeneous in high learning industries.

To summarize, we should expect a positive relationship between industry learning intensity and firm performance heterogeneity within an industry.

DATA AND EMPIRICAL ESTIMATION

Data

The data for this study come from two sources: Compustat and the US Census Bureau. There are two stages of analyses in this paper. First, we use *plant-level* data from the US Census Bureau to estimate the learning coefficients for each industry. We then use these estimated *industry* learning coefficients as independent variables in regressions that use Compustat data. These two data samples are described below.

First Stage Plant-level Sample (Census Bureau Data)

This sample is obtained from confidential micro-data available at the US Census Bureau. Since 1972, the US Census Bureau has been conducting a Census of Manufacturing (CM) every 5 years (there were two previous Censuses in 1963 and 1967). These censuses collect detailed *plant-level* data from all US manufacturing establishments with at least one employee. The data collected generally include the value of plant shipments, materials and energy inputs, employment, production hours, payments to labor, book values of physical assets, capital expenditures, inventories and ownership (single plant firm v. part of a multi-plant firm). In addition, the US Census Bureau also performs an Annual Survey of Manufactures (ASM) that collects similar data from a sample of US manufacturing establishments. In particular, the annual surveys are designed to get an overview of the sector during the inter-census years and hence, place considerable weight on large plants and plants belonging to multi plant firms. To account for new entrants, a sample of new entrants is added to the ASM sample every year.

The Census Bureau has collated the data from all these censuses and surveys and linked them through a longitudinal identifier to create a dataset (sometimes called the Longitudinal Research Database or LRD), which it makes available to researchers at Census Research Data Centers, subject to access restrictions and disclosure constraints. The most important disclosure constraint is that no data that can identify or relate to a single firm or plant

can be disclosed. Hence, in this paper, we do not identify statistics such as the median, minimum or maximum for variables obtained using Census Bureau data. For further details on this dataset, CM or ASM, please refer the US Census Bureau website.

Our sample is drawn from the LRD, which contains over 4 million plant-year observations from 1963 to 2001. Since the Census Bureau expends more effort on larger plants and firms, the quality of data for such cases is better. Also, they tend to have greater continuity of observations over time. In order to ensure reasonable data quality, we apply some sample selection criteria, the most important of which are:

- Eliminating all plants that were established before 1973 or after 1997. 1973 is the first year of the annual ASM and it is not possible to reliably obtain the entry year for plants that first appear in the 1963, 1967 or 1972 Censuses. In 1997, the Census Bureau switched from the SIC to NAICS classification system. Hence, we excluded plants established after 1997 to minimize errors from industry misclassifications.³
- Excluding all subsequent observations for a plant if the gap between two consecutive survey years for that plant is more than two years. This is done to ensure a higher reliability of our main variable, prior cumulative output.
- Removing all plants that have a Primary Industry Specialization Ratio (the output share of the primary 4-digit SIC industry in the case of a multi-product plant) of less than 75%. This is done to ensure homogeneity within an industry.
- Dropping outlier plants are in the top 0.5 percentile of capital-labor ratio or of growth in number of employees, shipments or capital expenditure.

The resulting sample contains 182,603 plant-year observations. Summary statistics for this sample are provided in Table 1a.

³ The older plants that continued after 1997 were assumed to stay in the same SIC-4 code as they were in 1997.

Second Stage Sample (Compustat Data)

We use Compustat to obtain data for the second stage sample. We compute firm q as the ratio of market value of assets to book value of assets, and firm profitability as the ratio of operating profits before depreciation to total assets.⁴ We then eliminate all outlying observations in the top and bottom 1% in terms of firm q or firm profitability. These data on firm performance are used to obtain the dispersion in firm q and firm profitability for each SIC-3 digit industry in each year from 1973 to 2000. We also obtained other industry level variables such as industry R&D and advertising intensity from Compustat. The industry classification was based on the primary industry code. The resulting sample contains 1,523 industry-year observations. Summary statistics for this sample are included in Table 1b.

Insert Table 1 about here

Variables

The important variables used in this study are described below. The first six pertain to the first stage plant-level sample and the last relates to the second stage industry-year sample.

Output. For any year before 1996, output for any plant is generally defined as the sum of the value of the plant's shipments (total plant revenues, deflated using SIC-4 industry-year deflators available on the NBER website) and the difference between year-beginning and year-

⁴ Specifically, the market value of assets is defined as the sum of the firm's market capitalization (data24*data25) and book value of assets (data6) less the book value of common stock (data60) and deferred taxes (data74). This definition follows Kaplan and Zingales (1997). Data13 in Compustat is taken to be the operating profit before depreciation.

ending deflated work in process and deflated finished goods inventories.⁵ For years including and after 1996, due to the unavailability of inventory data, output is simply defined to be the deflated shipments.

Value Added. Value Added is defined as the difference between real output and real materials (described below).

Labor. We define quantity of labor to be the labor hours expended in production worker equivalents. Labor hours for any plant are computed by dividing the total wage bill for the establishment by the average hourly wage for production workers in that establishment.

Materials. Real materials are defined as the sum of deflated cost of material purchases, external contract work, fuel and electricity.

Capital Stock and Capital Investment. We use the perpetual inventory approach to compute real capital stock. We compute separate stocks for buildings (or structures) and machinery. Real capital stock (k_{it}) in any given year, say for machinery, is computed as $k_{it} = (1 - d)k_{it-1} + I_{it-1} + R_{it}$ where d is an industry-year specific depreciation rate for machinery, I is the capital investment in machinery (deflated by an industry-year specific investment deflator for the year $t-1$) and R is the capitalized value of capital equipment rentals. If an establishment is not observed every year, we impute gross investment linearly (i.e. $I_{it} = 0.5 \cdot (I_{it} + I_{it-n}) \cdot (n-1)$, where I_{it} is the imputed investment for period t and n is the gap between the two survey years).

Prior Cumulative Output. This is used as a proxy for accumulated operating experience. Prior cumulative output is defined as the sum of real output through the end of the previous period i.e. $X_{it} = \sum(o_{i1}, o_{i2} \dots o_{it-1}) = X_{it-1} + o_{it-1}$, where o is real output. If an establishment is

⁵ This definition is identical to that implicitly used by the Census Bureau in its computations of plant “value added” (see below). The Census Bureau uses slightly different definitions in some industries due to differences in the nature of the manufacturing processes. We follow the Census Bureau’s definitions in all these cases. A detailed description is available on request.

not observed every year, we impute output linearly (i.e. $O_{it} = 0.5 * (o_{it} + o_{it-n}) * (n-1)$), where O_{it} is the imputed output for period t and n is the gap between the survey years).⁶

Heterogeneity of Firm Performance. We use firm q and firm profitability as measures of firm performance. Unlike most studies in the literature, we use direct measures of performance heterogeneity, specifically the cross-sectional dispersion of firm performance. We use three measures of cross-sectional dispersion of firm q and firm profitability. As the baseline measure, we take the difference between the 90th percentile and 10th percentile (of firm q or profitability) in an industry during a given year. The advantage of using this measure as opposed to say, variance is that it is an ordinal measure and hence, much less affected by the presence of outliers. As robustness checks, we use the inter-quartile range (or the difference between the 75th and 25th percentiles) and the standard deviation of firm profitability (or q) as other measures of heterogeneity.

Empirical Estimation

In the first part of our study, we use the Census data to estimate the learning rates for each SIC-3 industry and characterize the heterogeneity in industry learning rates. In the second part, which addresses the link between industry learning intensity and the heterogeneity of firm performance, we use the estimated industry-by-industry learning coefficients as explanatory variables in regressions with the range of firm performance as the dependent variable.

⁶ This measure of experience does not incorporate “organizational forgetting” and hence, does not differentiate between a small, old firm and a large, young firm. As rough robustness checks, we estimated (with OLS) the learning coefficients using (i) the cumulative output till $t-2$ as a measure of experience and (ii) including plant age as another variable in Equation 3. The learning coefficients so estimated were highly correlated with our baseline estimates

Measuring industry learning intensity. To proceed with empirical estimation of the importance of learning, we use the logarithmic version of Equation (2):

$$y_{ijt} = a_{jt} + \alpha_j \cdot k_{ijt} + \beta_j \cdot l_{ijt} + \lambda_j \cdot x_{ijt} + \varepsilon_{ijt} \quad (3)$$

Where

y is log (Value Added), k , l , and x are, log(L), log(K) and log(X) respectively

a is log(ϕ) and ε includes log(v)⁷. The choice of value added rather than output was primarily driven by the fact that some of the estimation approaches (described below) are applicable only to value added, and our desire to keep a consistent approach throughout the paper.⁸

The coefficient of interest is λ_j , the learning intensity of industry j .⁹ We estimate Equation (3) for *each* SIC-3 digit industry that has more than 50 plants. Estimating the production function industry-by-industry ensures that we are excluding the possibility that differences in returns to scale are being spuriously captured as learning. The terms a_{jt} in Equation (3) are coefficients on industry-year dummies, which capture all inter-temporal

⁷ There is, however, one subtle difference between Equations (2) and (3) in the interpretation of ε and a . In Equation (2), we interpreted ϕ and v to mean improvements in productivity resulting from factors other than learning-by-doing. In Equation 3, a and ε would also include classical measurement error in the dependent variable.

⁸ Nevertheless, where possible, we performed robustness checks using output as the dependent variable. The results closely match those with value added.

⁹ Since we treat the learning environment to be an industry characteristic, we estimate only one learning coefficient per industry. However, learning rates may change with time. As a robustness check, we estimated separate learning rates for the periods 1973-1984 and 1985-2000 (roughly equal sub-samples). The Spearman rank correlation between these two sets of coefficients was 0.54 and between these coefficients and our baseline estimates, 0.80 and 0.87 respectively, all statistically significant at or below the 0.01% level.

movements in the average industry productivity, including any industry-wide technology improvements. Hence, the econometric identification of the coefficients comes solely from cross-sectional deviations from the industry-year averages and not from changes in mean industry productivity over time.

First, we estimate equation (3) using OLS on the entire sample of plants. It is known that OLS estimates of labor are biased upwards while the capital coefficients are biased downward if there is unobserved plant specific heterogeneity (Griliches and Mairesse, 1995). Also, since including prior cumulative output on the right hand side is equivalent to adding a lagged dependent variable, the coefficient on prior cumulative output may be biased upward in the presence of persistent, unobserved heterogeneity. Another potential problem is that of survivor bias, which will bias the capital and prior cumulative output coefficients downward since firms with larger capital and experience can survive more negative “performance shocks”. To test for robustness to these concerns, we use three different estimation approaches in addition to OLS.

The first two approaches rely on straightforward extensions of the methods used in Olley and Pakes (1996) and in Akerberg, Caves and Frazer (2005), referred to as OP and ACF respectively in this paper. Here, we only sketch these methods. Appendix C provides some details as relevant to this paper. (For more technical discussions, please refer to the original articles.) Briefly, these methods rewrite equation 3 as

$$v_{ijt} = a_{jt} + \alpha_j \cdot k_{ijt} + \beta_j \cdot l_{ijt} + \lambda_j \cdot x_{ijt} + \omega_{ijt} + \theta_{ijt}$$

where ω_{ijt} is a firm-specific heterogeneity (e.g. management quality) that is not known to us but known to the firm and incorporated it into its decisions and θ_{ijt} is a completely exogenous error term. They then make two structural assumptions that (i) some observable decision (e.g. capital investments or materials) is an increasing function of ω , which lets us estimate the heterogeneity ω using the observed decision and that (ii) the decision on the level of some inputs at time ‘t’ is made at a time prior to ‘t’ (specifically, prior cumulative output or capital at

time 't' is completely determined at time t-1 and hence, are based only on ω_{ijt-1}), which then implies that these decisions must be unaffected by any changes between ω_{ijt} and ω_{ijt-1} . This lets us generate moment conditions that can be used in a GMM framework to estimate the coefficients. The two methods differ primarily on the assumptions about labor input (see Appendix C). These approaches address estimation problems due to endogeneity and survivor bias, albeit within bounds of their structural assumptions. The OP approach has been shown to generate unbiased productivity estimates and quite robust to a number of complications (Biesebroeck, 2004). The ACF approach being a simple extension of the OP approach will likely have similar properties. Nevertheless, arguments could be made that the underlying structural assumptions may not apply uniformly across all industries.

Finally, we use a specification with plant fixed effects. Simply adding plant fixed effects exacerbates some of the problems, particularly severely biasing the capital coefficient downward (Griliches and Mairesse, 1995). The same pattern holds for the experience coefficient, particularly with short panels. Nevertheless, we re-estimated equation (3) using plant fixed-effects on a "survivors sample" that excludes all plants that do not survive for at least 8 years.¹⁰ The disadvantages of this sample are the fewer industries with enough observations and the much larger standard errors of estimation.

To summarize, while all the above methods have their own potential shortcomings, our approach is to show that our subsequent results on performance heterogeneity are largely robust to these alternative estimation methods.

Inter-industry heterogeneity in learning. In addition to estimating individual learning coefficients for each industry, we also estimated a pooled regression to identify factors that may characterize inter-industry differences in learning rates. To do so, we adopted the following

¹⁰ The choice of 8 years, though a bit arbitrary, was driven by the fact that panels longer than 8 years meant much smaller sample sizes.

regression model, which includes a series of terms interacted with x_{ijt} , the cumulative output measure:

$$y_{ijt} = a_{jt} + \alpha \cdot k_{ijt} + \beta \cdot l_{ijt} + \lambda \cdot x_{ijt} + \lambda_1 \cdot C_{jt} \cdot x_{ijt} + \lambda_2 \cdot W_{jt} \cdot x_{ijt} + \lambda_3 \cdot R_{jt} \cdot x_{ijt} + \lambda_4 \cdot A_{jt} \cdot x_{ijt} + \varepsilon_{ijt} \quad (4)$$

Where

C is industry capital intensity (capital stock to employment ratio);

W is industry wages;

R is industry R&D intensity (R&D expenditure divided by sales); and

A is industry advertising intensity (advertising expenditure divided by sales).

As in the industry-specific learning regressions, the unit of analysis is plant-year, and we allow for industry-year dummies a_{jt} . However, we constrain the coefficients on capital and labor to be constant across industries (we tested the robustness of our results to this assumption by allowing these coefficients to vary by SIC-2 industry). For this analysis, industry R&D and advertising data are obtained from Compustat. We use OLS to estimate Equation 4 with plant fixed effects and instrumental variables specifications as robustness checks.

Industry learning intensity and heterogeneity of firm performance. In order to examine how industry learning is related to heterogeneity of firm performance, we use regressions of the following form:

$$\pi_{jt} = a_t + b \cdot \underline{\Delta}_t + c_1 \cdot R_{jt} + c_2 \cdot A_{jt} + c_3 \cdot C_{jt} + c_4 \cdot S_{jt} + c_5 \cdot N_{jt} + c_6 \cdot P_{jt} + \varepsilon_{jt} \quad (5)$$

Where π_{jt} is 90th to 10th percentile range of firm performance, either firm q or firm profitability, in industry j during year t ;

$\underline{\Delta}_t$ is the estimated industry learning intensity from Equation (3);

R is industry R&D intensity (R&D expenditure/sales);

A is industry advertising intensity (advertising expenditure/sales);

C is industry capital intensity (total assets/sales);

P is average industry profitability (operating profits/ total assets);

N is the number of firms in an industry; and,

S is industry size measured as total industry sales.

Note that the level of analysis here is the *industry-year*. For this analysis, we rely on data from Compustat. Hence, the only variable in equation (5) that comes from outside Compustat is the estimated learning intensity.

A brief note on the choice of control variables is in order. Intuitively, our earlier arguments on the inter-industry heterogeneity in learning rates also apply to any factor that increases complexity, tacitness or differentiation. R&D, advertising and capital intensity can logically be classified as such factors and would tend to increase the heterogeneity of firm performance. For instance, industries with high R&D or high advertising intensity may be very differentiated and hence, performance more dispersed. Furthermore, these are sunk costs, which increase the incentives for firms to stay in the market once they have incurred those costs (Gschwandtner and Lambson 2006), thereby increasing inter-firm heterogeneity. Average industry profitability may reflect inherent risk and hence, may be associated with a higher variance of returns. Finally, we add the number of firms and industry size as factors that may potentially increase measured heterogeneity.

RESULTS

Inter-Industry Heterogeneity In Learning-By-Doing

Table 2 presents the results of estimating Equation 3 for the *pooled* sample. Model 1 is a simple Cobb-Douglas production function, excluding the prior experience term that captures learning-by-doing. Model 2 expands on Model 1 by adding the prior experience term. The coefficient on prior cumulative output is 0.26, which implies a 16% gain in productivity for each doubling of cumulative output. This model, however, does not control for the possibility that the rate of technological improvement varies across industries. For instance, firms in an industry with significant technological advances may show productivity improvements even without learning-by-doing. A robust approach to address this is to include a dummy variable for each

industry-year combination, which will control for all inter-temporal changes (including technological improvements) in the average industry productivity. In other words, by including these fixed effects, we focus solely on deviations of plant productivities from industry-year averages. While our baseline definition of industry is at the SIC-3 digit level, the size of our pooled sample permits us to follow a far more conservative approach and use SIC-4 digit industry-year fixed effects. Model 3 includes 9,967 *separate* SIC-4 industry-year dummies, which control for all technological improvements in each SIC-4 industry (and consequently, in each SIC-3 industry). The estimated learning coefficient falls to 0.23 when these controls are added.

We then estimated Equation 3 using OLS for each of the 117 SIC-3 industries that has more than 50 plants. Models 4-1 to 4-117 allow each industry to have its own coefficients on capital, labor, and prior cumulative output. They also allow for year dummies within each SIC-3 industry and hence, control for SIC-3 industry-wide productivity improvements. Given space constraints, we present only the coefficients on cumulative output from these models in Appendix A (Column 3). Coefficients on capital and labor are available on request. Figure 1 presents a histogram of the coefficients on cumulative output. As expected, there is a significant variation in learning intensities across industries, ranging from just above 0 to almost 0.60 with an average of 0.22 (almost identical to the estimate for the pooled sample in Model 3).¹¹ We also estimated Equation 3 using a plant-fixed effects specification for a smaller sample (49 industries) that includes only firms that survive for at least 8 years. Figure 1b presents a histogram of these learning coefficients. The distribution of the learning coefficients is similar for the OLS and fixed effects specifications though the mean coefficient using the fixed effects

¹¹ When output is used as the dependent variable (instead of value added), the average learning intensity across the 117 industries is about 0.11. Learning coefficients based on output as a dependent variable are available on request.

specification is lower (as expected).

 Insert Figure1 about here

Finally, we used the OP and ACF approaches explained above to estimate the learning rates for industries with sufficient data. Appendix A presents the industry-by-industry learning coefficients from Equation 3 using the four different approaches, as well as the industry ranks based on the OLS learning coefficient (Column 2). While the individual coefficients vary in magnitude across the different methods, they are highly correlated with each other. The Spearman's rank correlations among OLS, OP and ACF estimates range from 0.57 to 0.74 and are statistically significant at the 0.01% level. The estimates from fixed effects specification with value added as a dependent variable are, however, much less correlated with the others (0.09 to 0.18), and the correlations are statistically insignificant.¹²

 Insert Table 2 about here

We now try to characterize the heterogeneity in learning intensity. From Appendix A, we can see that the top 6 industries based on the OLS learning coefficient are 357 (computers), 283 (pharmaceuticals), 291 (petroleum refining), 386 (photographic equipment & supply), 287 (agricultural chemicals) and 289 (miscellaneous chemicals). The lowest in the list are 317 (leather goods) 322 (glass products), 262 (paper mills), 228 (yarn & thread mills) and 311

¹² With output as a dependent variable for the fixed effects specification, the correlation improves considerably to about 0.28 and becomes statistically significant at the 5% level. For purposes of consistency, however, we work with value added as a dependent variable.

(leather tanning). This list suggests a positive association between complex, knowledge-intensive and capital-intensive settings, and the rate of learning-by-doing.

To test this more formally, we estimate Equation 4 using OLS. Again, our pooled sample permits us to adopt a more conservative approach and use a finer SIC-4 digit industry definition. The variables of interest are the interaction terms between prior experience and industry factors. Models 5 and 6 in Table 3 use a larger sample for which we have complete data on industry wages and the capital-labor ratio, omitting the industry R&D and advertising intensity terms. Model 5 includes only year indicators while Model 6 includes a full set of industry-year dummies. In both cases, the learning coefficient is significantly higher in industries with greater capital intensity. The interaction effect of industry wages on prior experience becomes insignificant once industry-year effects are controlled for.

Model 7 estimates Equation (4) with a smaller data set for which we have complete industry R&D and advertising data from Compustat. The coefficient on prior experience is significantly higher in industries with higher capital-labor ratio and greater R&D and advertising intensity. Models 5-7 assume that the coefficients on capital and labor are the same across industries; Model 8 repeats the tests in Model 7, allowing the coefficients on capital and labor to vary by SIC-2 digit industry.¹³ The results in Model 8 are not substantially different from Model 7. Finally, we estimate Models 9 and 10 to check if these results are robust to the inclusion of plant fixed effects. Model 9 does not include any of the direct terms while Model 10 includes them. While the effects decrease considerably in magnitude (as expected) in Model 9, the direction and statistical significance of the results persist. The significance of the interaction terms increase considerably when the direct terms are included. However, we must be cautious because the year-on-year changes in variables such as R&D intensity are very small. Finally,

¹³ The choice of SIC-2 industry was dictated by the cost of computation. Allowing the capital and labor coefficient to vary by SIC-3 industry would have meant estimating over 350 coefficients.

we used an instrumental variables specification with once-lagged variables as instruments. The economic substance and statistical significance of these results (available on request) were very similar to those in Models 7 and 8.

To summarize, we find that the learning rate increases with industry capital intensity, R&D intensity, and advertising intensity, factors that are likely to reflect process complexity, knowledge tacitness and product differentiation.

 Insert Table 3 about here

Industry Learning Intensity And Firm Performance Heterogeneity.

We now examine the important question of how these variations in industry learning intensity are related to firm performance heterogeneity. Figure 2 gives an idea of the potential linkages. It presents the distributions of firm profitability and firm q (relative to the industry-year average) for “high learning” (i.e. industries with learning rates above the median learning rate) and “low learning” industries. Figure 2 shows that in the “high learning” industries the profit distribution has a thicker left tail; and both tails of the firm q distribution are thicker. Thus, both measures of firm performance have greater dispersion in “high learning” industries. To test this formally, we use the industry estimated learning coefficients (shown in Appendix A) as independent variables in Equation 5, with the range of firm performance within an industry as the dependent variable.

Table 4 presents the results of estimating Equation 5 for our two measures of performance heterogeneity. Since all the variables except the learning coefficient have very skewed distributions, we use their logarithms rather than the original values. Models 11 to 14 use the range of firm profitability as the dependent variable while Models 15 to 18 use the range of firm q . The results are strongest for heterogeneity in firm profitability (Models 11 to 14). The

“industry learning intensity” coefficient varies from 0.203 with learning estimates using fixed effects specifications to 0.926 when OLS based learning estimates are used. However, regardless of the approach adopted, the coefficient on industry learning intensity is always statistically significant. This implies that the difference between the “best performers” and the “worst performers” is considerably greater in industries with high learning. Similar results hold when we use firm q as a measure of firm performance. With OLS, OP and ACF based learning coefficients (Models 15-17), the coefficient on industry learning is positive and statistically significant at the 1% level. With plant fixed-effects based learning coefficients (which are available only for a smaller set of industries), the coefficient turns statistically insignificant. However, we must note that the learning coefficients themselves are estimated with error and with greater error in non-OLS specifications; this implies the estimated coefficients on industry learning intensity in Table 4 are *downward-biased* and more so when using learning estimates from non-OLS specifications. In summary, it appears reasonable to conclude that firm performance-heterogeneity is greater in industries with higher learning intensities.

Turning to the other coefficients in the regressions, sunk costs (particularly R&D and advertising, and to an extent, capital intensity) are also linked to increased dispersion of firm performance. This is broadly in line with Gschwandtner and Lambson (2006), who found that sunk costs tend to increase profit variability in an industry, and theoretical industry models in the economics literature such as Hopenhayn (1992) that predict increased productivity dispersion due to higher sunk costs.¹⁴ The mean industry profitability appears to be associated with increased heterogeneity of firm value (which is consistent with a higher risk-higher return story) but rather counterintuitively, with decreased profitability dispersion. Another counterintuitive result is that large industries (by industry sales) tend to have lower heterogeneity. We do not

¹⁴ In an unpublished working paper, Rivkin (2001) also finds that the dispersion of firm profit rates is higher in industries with opportunities for R&D and product differentiation.

have a good explanation for this except, perhaps, that they may be mature industries. Industries with many firms, in line with our intuition, show a wider dispersion.

 Insert Table 4 about here

Robustness Checks

We performed a series of tests to confirm that we are most likely measuring the effect of learning-by-doing and that our subsequent results on the heterogeneity of firm performance are robust to alternative specifications. Briefly, the tests show that factors such as survivor bias, sample selection, R&D investments, measurement errors in capital, choice of production function form, and industry life cycle effects are not driving the observed heterogeneity in learning rates. The details are presented in Appendix B. Based on our tests, it appears safe to conclude that our estimates provide a reasonably robust rank ordering of industries based on their learning rates.

We tested the robustness of our results on the connection between learning and firm performance by running the same type of regressions as in Table 4, but with different measures of performance heterogeneity, level of aggregation, choice of time periods, assumptions about error correlation structures etc. Table 5 presents the coefficients on the learning estimates from those regressions. (Each line in Table 5 is comparable to line 1 from Table 4). In a majority of the regressions, learning shows a significant positive association with the heterogeneity of firm performance. Appendix B presents additional details.

DISCUSSION AND CONCLUSION

It is widely accepted and documented in the strategy literature that the industry environment affects competitive heterogeneity. In addition, many studies have shown that the rate of organizational learning varies greatly across firms and industries. Nevertheless, the

connections between learning intensity and the industry competitive environment have not been systematically explored, given data limitations. Our paper provides quantitative evidence that supports the case for treating learning intensity as a fundamental characteristic of the industry, much as R & D and advertising. Furthermore, our paper facilitates empirical implementations of such a concept by providing reasonably comparable estimates of learning intensity for a wide range of industries, encompassing most of the US manufacturing sector.

To our knowledge, this study is the first to provide a quantitative comparison of learning rates across such a broad set of industries. The range of learning slopes from the OLS learning coefficients (computed as $2^{-\lambda}$, where λ is the learning coefficient) varies from close to 100% (or no learning) in SIC 262 (Paper Mills) to a maximum of 68% in SIC 283 (Drugs) and 66% in SIC 357 (Computers).¹⁵ This is comparable to the range of estimates in prior studies. For instance, the survey by Dutton and Thomas (1984) found that the median learning slope in about 22 industry-specific studies was about 80% and the range was from 55% to 110%. In our analysis, the range of learning slopes obtained from other estimation approaches differ, but not unreasonably so (OP: 57% to 106%; ACF: 51% to 107%; FE: 63% to 112%). All are consistent with the range of learning slopes found in prior work.

Our study goes beyond simply establishing the heterogeneity in learning rates to identify some broad patterns in these learning rates, as indicated in Table 3. Even within the limited interpretation permitted by our crude proxies, the results are consistent with the argument that learning from own experience may be more important in environments where process complexity is high, tacit knowledge is critical, and products are differentiated. Knowledge transfers between firms, and perhaps even within firms, is naturally harder in such environments, and firms may have to rely more on their experience. Though intuitive, this study

¹⁵ A learning slope of $x\%$ implies that a doubling of cumulative output leads to a $(100-x)\%$ increase in productivity.

is the first attempt to quantify these patterns in a systematic way across a broad sample of industries.

The second contribution of our paper is to demonstrate a robust association between the industry learning intensity and the cross-sectional heterogeneity of firm performance. The results in Tables 4 and Table 5 show that firm performance is more heterogeneous in high learning industries. More importantly, the economic significance of this effect seems to be large. Based on the coefficients in Model 11, an increase of one standard deviation in the learning coefficient (0.097) is associated with a 31% ($=0.926 \times 0.097 / 0.2891$) increase in the profitability range. Similar calculations with Models 12, 13 and 14 give estimates of 24%, 29% and 10% respectively. These are comparable to or even higher than the effect of an increase in R&D or advertising intensity. For instance, using coefficients from Model 11, an increase in R&D intensity by one standard deviation from the mean (i.e. from 0.0217 to 0.0443) is associated with a 12% ($=0.050 \times [\ln(0.0443) - \ln(0.0217)] / 0.2891$) increase in profitability range.¹⁶

Thus, we have shown that the cross-sectional variation in firm performance increases with the learning intensity of the industry. While a detailed empirical examination of the underlying reasons is beyond the scope of this paper, we briefly offer some “informed conjectures”. Regressions similar to those in Table 4 with the top and bottom deciles of the firm profit distribution as the dependent variable show that the increased dispersion in firm profits is almost entirely due to a *downward shift* in the 10th percentile of profits rather than an upward shift in the 90th percentile. In contrast, the heterogeneity of firm q appears to be mostly due to an upward shift in the 90th percentile rather than a decrease in the 10th percentile. These results

¹⁶ The calculated impact of learning may also be partly high because it is measured with greater error than R&D (and hence, a standard deviation increase in learning may be “more” than a standard deviation increase in R&D). However, the fact that all approaches give the same order of magnitude should give us some confidence in the assessment.

are consistent with a story where learning is uncertain and only the “learning the right thing” is rewarded. Additional support for this argument comes from examining the link between learning intensity and the mean firm performance in an industry. Specifically, if uncertain learning creates “winners” that grow to be bigger than others, the *size-weighted average* firm performance (e.g. total industry profits divided by total industry assets) should be significantly higher in industries with high learning. Indeed, this industry profit ratio is about 14% in “high learning” industries compared to 11% in “low learning” industries. The size-weighted industry average q is 1.29 in “high learning” industries, vs. 1.08 in “low learning” industries. Moreover, regressions similar to those in Table 4 with the mean industry profitability as the dependent variable show a strong positive relation between learning intensity and mean industry profitability. Though by no means conclusive, such evidence points to a story more like the uncertain learning-by-doing in Levitt and March (1988) and Levinthal and March (1991) rather than the sure-shot learning curve often assumed in the economics literature.

As with all empirical studies, our analysis comes with a number of limitations. We adopt a highly aggregated view of learning by focusing on learning intensity at the industry level. Clearly, there is likely to be considerable heterogeneity in products and learning rates within industries, perhaps even greater than the inter-industry variations. Moreover, our study does not shed any light on the mechanisms of learning e.g. factors within organizations such as training and engineering activities (Adler and Clark, 1991), structures and routines (Nelson and Winter, 1982) that may affect learning. Furthermore, it is to be expected that the meaning and context of organizational learning vary significantly across (and within) industries. Since this paper follows a purely empirical approach and infers the importance of learning by examining the coefficient on cumulative output, the findings from this study are necessarily a very simplified and stylized representation of the learning environment. Furthermore, learning-by-doing is only one form of organizational learning (Levitt and March 1988; Malerba 1992). There are many

other forms of learning such as learning from others, which are not examined in this study. Finally, there are measurement issues that commonly afflict studies of productivity estimation. Notwithstanding these limitations, we believe that this aggregate approach provides a “big picture” view of the heterogeneity in industry learning-environments that complements detailed micro-level studies of learning.

Although not a limitation, per se, the interpretation of the learning coefficient deserves some discussion. As we measure it, the learning coefficient does not reveal two aspects of learning that have been considered in prior studies. First, it does not tell anything about spillovers of learning across firms. Firms can apply experience gained from one product to cost reduction or quality improvements in other (perhaps, similar) products (Argote, 1999; Benkard, 2001; Irwin and Klenow, 1994) and it is reasonable to expect that learning “spills over” from one firm to another. In our approach, all improvements resulting from such industry-wide learning spillovers are captured by the industry-year dummy variables. As an extreme, an industry where some firms learn considerably from experience but *all* knowledge so generated is transferred to other firms *immediately* (thereby leaving the relative performances unchanged) would be measured as having a zero rate of learning. However, in such a case, learning rates would not affect firm performance and a “zero” learning coefficient would not entirely be meaningless. The second issue is organizational forgetting. It has been established that the knowledge accumulated through learning depreciates rapidly (Argote, 1999; Benkard, 2001). The use of a single learning coefficient clearly masks underlying differences in the rate of depreciation. A low learning coefficient could mean either a low learning rate combined with low rate of forgetting, or a high learning rate combined with rapid depreciation. Unfortunately, our data do not permit us to reliably estimate the rates of forgetting. Even so, the learning coefficient can still be meaningfully interpreted as “net rate of learning” or as the “net importance” of experience in the production process. Moreover, there is almost nothing in the

prior literature that discusses whether discriminating between these scenarios has any managerial or policy implications. Nevertheless, because this approach suppresses all differences in spillovers and depreciation, we must be cautious when referring to “high learning” or “low learning” industries. They may not always be industries where firms “learn the most” or “learn the least”; rather they refer to industries where the firm’s *own* experience is more or less important in the production process.

The present study establishes basic relationships between industry learning and firm performance, but many extensions are possible. One would be to examine the mechanisms that explain the link between learning intensity and the heterogeneity of firm performance. We have provided some potential theoretical reasons for this association; however, their importance must be sorted out. More broadly, the learning estimates from this study can be used to analyze how variations in industry learning rates affect firm behavior. Strategic choices that may be affected by learning include incentive structures, governance structures, investments in innovation, capital and technology, and perhaps even organizational structures and processes. For instance, a bigger role for learning from experience may require an incentive structure oriented towards long-term performance goals rather than short-term ones. Another interesting line of inquiry would be to generalize our findings to examine how variations in the knowledge acquisition processes (e.g. through own learning vs. inter-firm spillovers vs. intra-firm spillovers etc.) across industries affect the observed heterogeneity.

Heretofore, researchers have been constrained by the non-availability of learning rate data across a wide range of industries needed to address such questions. This study provides industry specific estimates that may be used to explore the role of learning by doing, particularly in broader inter-industry contexts. We must, however, reemphasize that we do not claim to provide a single “correct” set of learning rates. Rather, we offer learning rates from different estimation methods that used in combination, provide a reasonably robust rank ordering of

industries based on their learning rates.

REFERENCES

- Akerberg, D., Caves, K. and Frazer, G. Structural estimation of production functions: An application to the timing of input choices, mimeo, August 2005
- Adler, P.S. & Clark, K.B. 1991. Behind the learning curve: A sketch of the learning process. *Management Science*. 37: 267-281.
- Argote L. 1999. *Organizational learning: creating, retaining and transferring knowledge*, Kluwer Academic Publishers.
- Bahk, B.H. & Gort M. 1993. Decomposing learning-by-doing in new plants, *The Journal of Political Economy*, 101: 561-583.
- Barney, J. B. 1991. Firm resources and sustained competitive advantage, *Journal of Management*, 17: 99-120
- Baum, J.A.C. & Ingram, P. 1998. Survival-enhancing learning in the Manhattan hotel industry, 1898-1980. *Management Science*, 44: 996-1016.
- Benkard, C.L. Learning and forgetting: The dynamics of aircraft production. *American Economic Review*, 90: 1034-1054.
- Cohen, W.M. & Levinthal, D.A. 1989 Innovation and learning: the two faces of R&D. *Economic Journal*, 99: 569-596.
- Cohen, W.M. & Levinthal, D.A. 1990 Absorptive capacity – a new perspective on learning and innovation. *Administrative Science Quarterly*, 35: 128-152.
- Dierickx, I. & Cool, L. 1989. Asset accumulation and sustainability of competitive advantage. *Management Science*, 35: 1504-1511
- Dutton, J.M. & Thomas, A. 1984. Treating progress functions as a managerial opportunity. *Academy of Management Review*, 9: 235-247.
- Griliches Z. & Mairesse, J. 1995. *Production functions: the search for identification*. Working paper no. 5067, National Bureau of Economic Research.
- Gruber, H. 1994. The Yield Factor and the Learning Curve in Semiconductors. *Applied*

- Economics*, 26: 225-248.
- Gruber, H. 2000. The Evolution of Market Structure in Semiconductors: The Role of Product Standards. *Research Policy* 29: 725-740.
- Gschwandtner, A. & Lambson, V.E. 2006. Sunk costs, profit variability and turnover. *Economic Inquiry*, 44(2): 367-373
- Hoopes, D.G., Madsen, T.L & Walker, G. 2003. Guest editors' introduction to the special issue: Why is there a resource-based view? Toward a theory of competitive heterogeneity. *Strategic Management Journal*, 24: 889-902
- Hopenhayn, H. 1992. Entry, exit and firm dynamics in the long run. *Econometrica*, 60(5): 1127-1150
- Hayes, R.H. & Clark, K.B. 1986 Why some factories are more productive than others. *Harvard Business Review*, 64(5): 66-73.
- Irwin, D.A. & Klenow, P.J. 1994. Learning-by-doing spillovers in the semiconductor industry. *The Journal of Political Economy*, 102: 1200-1227.
- Kaplan, S.N. & Zingales, L. 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics*, 112(1): 169-215
- Levinthal, D.A. & March, J.G. 1993. The myopia of learning. *Strategic Management Journal*, 14: 95-112.
- Levitt, B. & March, J.G. 1988. Organizational learning. *Annual Review of Sociology*. 14: 319-340
- Lieberman, M.B. 1984. The learning curve and pricing in the chemical processing industries. *Rand Journal of Economics*, 15(2): 213-228
- Lieberman, M.B. 1987. The learning-curve, diffusion, and competitive strategy. *Strategic Management Journal*. 8: 441-452.
- Lippman, S.A. & Rumelt, R.P. 1982 Uncertain imitability - an analysis of interfirm differences in efficiency under competition. *Bell Journal of Economics*, 13: 418-438.

- Malerba, F. 1992. Learning by firms and incremental technical change. *Economic Journal*, 102: 845-859.
- Muth, J.F. 1986. Search theory and the manufacturing progress function. *Management Science*, 32: 948-962
- Nelson, R.R. & Winter, S.G. *An evolutionary theory of economic change*. Harvard University Press. Cambridge, MA.
- Olley, G.S. & Pakes, A. 1996. The dynamics of productivity in the telecommunication equipment industry. *Econometrica*, 64(6): 1263-1297
- Peteraf, M.A. 1993. The cornerstones of competitive advantage: a resource based view. *Strategic Management Journal*, 14(3): 179-191
- Pisano G.P., Bohmer R.M.J. & Edmondson A.C. 2001. Organizational differences in rates of learning: Evidence from the adoption of minimally invasive cardiac surgery. *Management Science*, 47: 752-768
- Rapping, L. 1965. Learning and World War II production functions. *Review of Economics and Statistics*, 47: 81-86.
- Rivkin, J.W. 2000. Imitation of complex strategies. *Management Science*, 46(6): 824-844
- Rivkin, J.W. 2001. Reconcilable Differences: The Relationship Between Industry Conditions and Firm Effects, *Unpublished Doctoral Dissertation*, Harvard Business School
- Rumelt, R.P. 1991. How much industry matter? *Strategic Management Journal*, 12: 167-185
- Schmalensee, R. Inter-industry studies of structure and performance. In R.Schmalensee and R.d. Willig (Eds.), *Handbook of Industrial Organization*, Vol II: 952-1009.
- Sinclair, G., Klepper, S. & Cohen, W. 2000. What's experience got to do with it? Sources of cost reduction in a large specialty chemicals producer. *Management Science*. 46: 28-45.
- Spence, A.M. 1981 The learning curve and competition. *Bell Journal of Economics*, 12: 49-70
- Thornton, R.A. & Thompson, P. 2001. Learning from experience and learning from others: An exploration of learning and spillovers in wartime shipbuilding, *American Economic Review*, 91:1350-1368.
- von Hippel, E. 1988. *The Sources of Innovation*. New York: Oxford University Press.

Figure 1: Inter-industry Heterogeneity in Learning-by-Doing

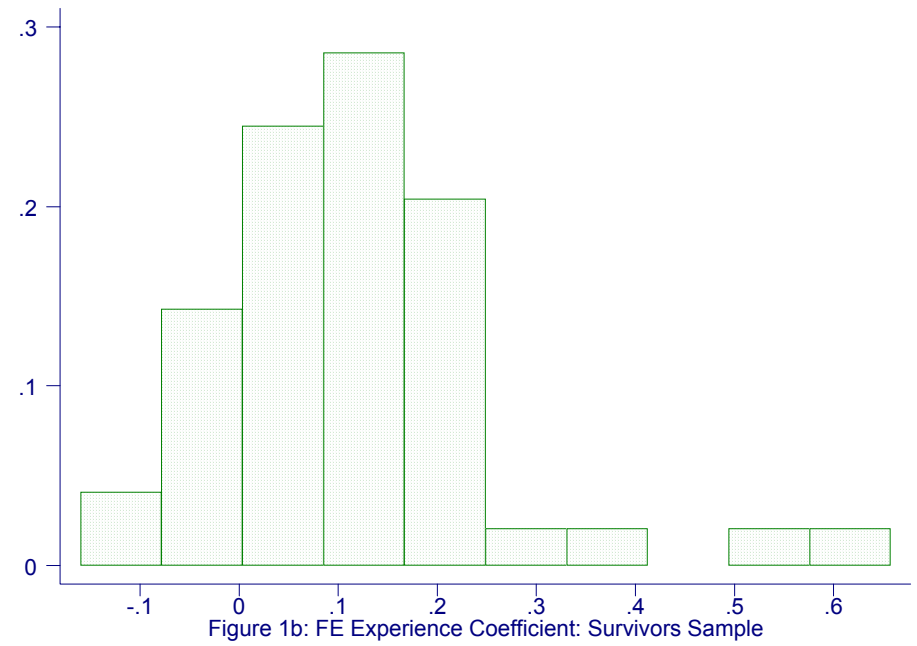
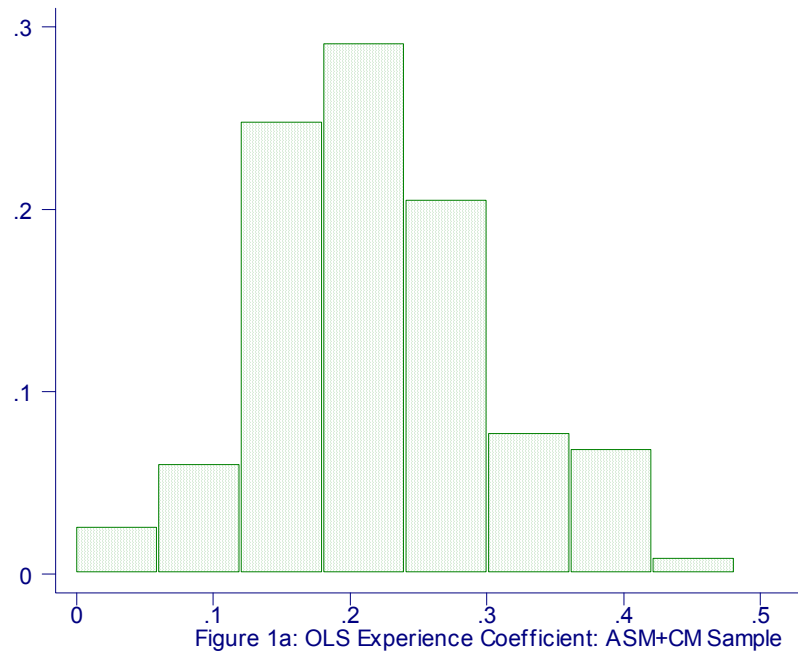


Figure 2: Industry Learning Environments and Firm Performance Heterogeneity

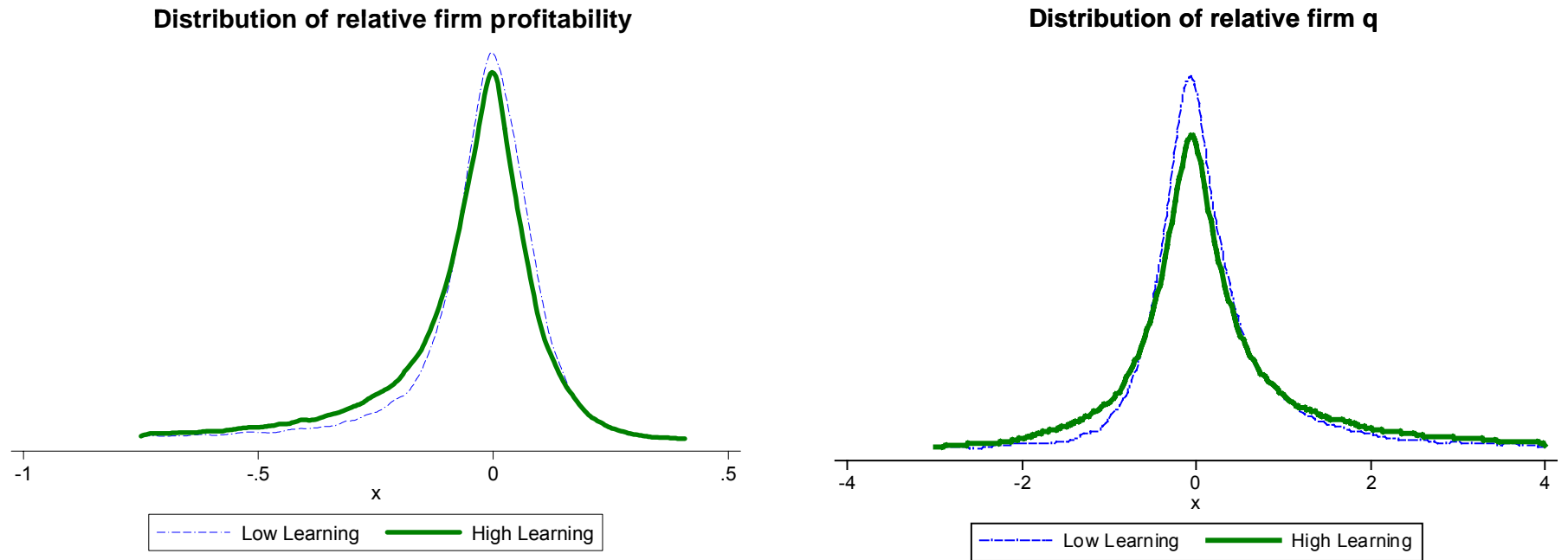


TABLE 1a
Overall Descriptive Statistics (Census or First Stage Sample)^a

| Variable | Mean | s.d | Min | Max | 1 | 2 | 3 | 4 | 5 |
|-----------------------------|------|------|-----|-----|--------|--------|--------|--------|--------|
| 1. Output ^{b/c} | 8.20 | 1.85 | | | | | | | |
| 2. Value Added ^c | 7.40 | 1.86 | | | .93*** | | | | |
| 3. Capital | 6.62 | 2.12 | | | .81*** | .77*** | | | |
| 4. Labor | 4.59 | 1.59 | | | .87*** | .85*** | .75*** | | |
| 5. Material | 7.29 | 2.12 | | | .94*** | .81*** | .77*** | .79*** | |
| 6. Prior Experience | 9.21 | 2.24 | | | .89*** | .83*** | .83*** | .80*** | .85*** |

Table 1b
Overall Descriptive Statistics (Compustat or Second Stage Sample)

| | | | | | | | | | |
|--|-------|--------|-------|-------|----------|----------|---------|-------|--------|
| 1. Industry R&D intensity | 0.02 | 0.02 | 0.00 | 0.12 | | | | | |
| 2. Industry advertising intensity | 0.02 | 0.02 | 0.00 | 0.08 | -0.08*** | | | | |
| 3. Number of firms | 44.15 | 51.31 | 10 | 453 | 0.72*** | -0.09*** | | | |
| 4. Industry sales (\$ billion) | 47.59 | 140.40 | 0.44 | 1,498 | 0.12*** | -0.10*** | 0.26*** | | |
| 5. Industry q range (10 th -90 th pctile) ^c | 1.65 | 1.36 | 0.093 | 10.86 | 0.47*** | 0.16*** | 0.41*** | 0.03 | |
| 6. Industry profitability range (10 th -90 th) ^c | 0.28 | 0.16 | 0.03 | 1.59 | 0.39*** | 0.07*** | 0.38*** | -0.01 | 0.63** |

a. There are two separate samples. Table 1a refers to the first stage plant-level sample ($n = 182,603$ of which 170,666 are in SIC-3 industries that have at least 50 plants) for which we are not able to present the minimum and maximum due to disclosure restrictions on data pertaining to individual firms. Table 1b refers to the second stage sample based on Compustat data ($n=1,523$)

b. Variables 1-6 in Table 1a are logarithms of their original values. Please refer text for precise definition of variables.

c. These variables are always *dependent* variables, hence, correlations between them are not meaningful.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

TABLE 2

Pooled Learning Coefficients

| Variable | | Model 1 (OLS) | Model 2 (OLS) | Model 3 (OLS) | Models 4-1 to 4-117(OLS) |
|-------------------------|------------------|---------------------|--------------------|------------------------|--|
| 1. | Capital | 0.271*** (0.003) | 0.12*** (0.003) | 0.07*** (0.003) | -0.12 to 0.29 (Details available on request) |
| 2. | Labor | 0.710*** (0.004) | 0.59*** (0.004) | 0.66*** (0.004) | 0.34 to 0.98 (Details available on request) |
| 3. | Prior Experience | | 0.26*** (0.004) | 0.23*** (0.004) | 0.00 to 0.60 (Details available in Appendix A) |
| 4. | Fixed Effects | Year | Year | SIC-4 Industry-Year | SIC-3 Industry- Year |
| N | | 213,256 | 170,666 | 170,666 | 83 to 7,244 (Details available on request) |
| R ² | | 0.80 | 0.81 | 0.85 | |
| Adjusted R ² | | 0.80 | 0.81 | 0.85 | |

a. The unit of analysis is *plant-year*. Value added is the dependent variable. Coefficients on dummies not presented.

b. Models 4-1 to 4-117 are 117 separate OLS estimations along the lines of Model 2, one for each SIC-3 industry.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Standard errors in parentheses (Models 1-3 present clustered standard errors that allow for arbitrary autocorrelation of errors within a plant).

TABLE 3: Inter-industry Heterogeneity in Learning-by-doing^a

| | Model 5 (OLS) | Model 6 (OLS) | Model 7 (OLS) | Model 8 ^b (OLS) | Model 9 ^b (FE) | Model 10 ^b (FE) |
|---|----------------------|---------------------|---------------------|-------------------------------|------------------------------|-------------------------------|
| 1. Capital | 0.064*** (0.003) | 0.074*** (0.003) | 0.077*** (0.005) | By SIC-2 | By SIC-2 | By SIC-2 |
| 2. Labor | 0.647*** (0.004) | 0.665*** (0.004) | 0.667*** (0.007) | By SIC-2 | By SIC-2 | By SIC-2 |
| 3. Prior Experience | 0.250*** (0.01) | 0.225*** (0.01) | 0.198*** (0.01) | 0.191*** (0.01) | 0.066*** (0.01) | 0.022* (0.01) |
| 4. Experience*Industry Wages ^c | -1.14*** (0.31) | -0.526 (0.37) | 0.083 (0.52) | 0.069 (0.57) | 0.742*** (0.14) | 0.036 (0.475) |
| 5. Experience*Industry Capital Labor Ratio ^c | 0.428*** (0.05) | 0.323*** (0.32) | 0.226*** (0.06) | 0.266*** (0.07) | 0.049*** (0.01) | 0.424*** (0.04) |
| 6. Experience*Industry Advertising Intensity | | | 0.869*** (0.14) | 0.946*** (0.16) | 0.156*** (0.03) | 0.553*** (0.14) |
| 7. Experience*Industry R&D Intensity | | | 0.446*** (0.10) | 0.503*** (0.11) | 0.037*** (0.01) | 1.35*** (0.08) |
| 8. Industry Wages | 0.044*** (0.003) | NA | NA | NA | | 0.004 (0.01) |
| 9. Industry Capital Labor Ratio ^c | -0.036*** (0.004) | NA | NA | NA | | -0.446*** (0.04) |
| 10. Industry R&D Intensity ^d | | NA | NA | NA | | -0.149*** (0.01) |
| 11. Industry Advertising Intensity ^d | | NA | NA | NA | | -0.047*** (0.02) |
| 12. Fixed Effects | Year | Industry-Year | Industry-Year | Industry-Year | Plant & Year | Plant & Year |
| N | 182,603 | 182,603 | 71,824 | 71,824 | 71,824 | 71,824 |
| R ² | 0.82 | 0.83 | 0.83 | 0.83 | 0.91 | 0.93 |

a. The unit of analysis is plant-year. Log value added is the dependent variable. Coefficients on dummies not presented. With industry-year effects, direct terms are not included because the dummies control for all changes at the industry-year level (e.g. industry R&D intensity).

b. Models 8 to 10 allow for *industry-specific* (by SIC-2) capital and labor coefficients. Given the large number of capital and labor coefficients, we do not present them here.

c. Coefficients and standard errors divided by 1000 for presentation purposes.

d. Coefficients and standard errors divided by 100 for presentation purposes.

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Standard errors in parentheses (Models 5-8 present clustered standard errors, which allow for arbitrary autocorrelation within a plant).

TABLE 4

Learning-by-doing and Heterogeneity in Firm Performance^a

| | Model 11 | Model 12 | Model 13 | Model 14 | Model 15 | Model 16 | Model 17 | Model 18 |
|---|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Dependent Variable | Profit Dispersion | Profit Dispersion | Profit Dispersion | Profit Dispersion | q dispersion | q dispersion | q dispersion | q dispersion |
| Method of estimating learning coefficient | OLS | OP | ACF | FE-SURV | OLS | OP | ACF | FE-SURV |
| 1. Industry Learning Intensity | 0.926*** (0.10) | 0.485*** (0.09) | 0.491*** (0.06) | 0.203** (0.08) | 1.41*** (0.14) | 0.803*** (0.12) | 0.677*** (0.08) | 0.027 (0.13) |
| 2. Industry R&D Intensity | 0.050*** (0.01) | 0.048*** (0.01) | 0.046*** (0.01) | 0.049*** (0.01) | 0.102*** (0.01) | 0.097*** (0.01) | 0.097*** (0.01) | 0.075*** (0.02) |
| 3. Industry Advertising Intensity | 0.040*** (0.01) | 0.040*** (0.01) | 0.035*** (0.01) | 0.030*** (0.01) | 0.070*** (0.01) | 0.068*** (0.01) | 0.062*** (0.01) | 0.058*** (0.01) |
| 4. Industry Profitability (Mean) | -0.131*** (0.04) | -0.075* (0.04) | -0.101** (0.04) | -0.044 (0.05) | 0.207*** (0.05) | 0.286*** (0.05) | 0.265*** (0.05) | 0.345*** (0.06) |
| 5. Industry Capital Intensity | -0.110** (0.05) | -0.047 (0.05) | -0.054 (0.05) | -0.095* (0.05) | 0.113* (0.06) | -0.172** (0.07) | 0.114** (0.06) | 0.144** (0.07) |
| 6. Industry Sales | -0.117*** (0.01) | -0.105*** (0.01) | -0.125*** (0.01) | -0.100*** (0.01) | -0.117*** (0.01) | -0.100*** (0.01) | -0.126*** (0.01) | -0.095*** (0.02) |
| 7. Industry Number of Firms | 0.251*** (0.02) | 0.240*** (0.02) | 0.256*** (0.02) | 0.244*** (0.02) | 0.268*** (0.02) | 0.249*** (0.02) | 0.277*** (0.02) | 0.311*** (0.03) |
| 8. Fixed Effects | Year | Year | Year | Year | Year | Year | Year | Year |
| N | 1,523 | 1,523 | 1,523 | 979 | 1,523 | 1,523 | 1,523 | 979 |
| R ² | 0.38 | 0.37 | 0.37 | 0.39 | 0.55 | 0.54 | 0.54 | 0.55 |
| Adjusted R ² | 0.37 | 0.35 | 0.36 | 0.37 | 0.54 | 0.53 | 0.53 | 0.53 |

a. The unit of analysis in all these regressions is SIC-3 industry-year. All variables except the industry learning intensity are logarithms of their original values. The dependent variable is the (logarithm of) difference between 90th and 10th percentiles of either firm profitability (Models 11-14) or firm q (Models 15-18). Coefficients on dummies not presented.

b. Models 11 and 15 use learning intensities computed using OLS. Models 12 and 16 use Olley-Pakes (OP) based learning coefficients while Models 13 and 17 use an extension of the approach outlined in Akerberg et al (2005). Models 14 and 18 use learning intensities computed using plant fixed effects on the sample of survivors (at least 8 years).

* $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Robust standard errors in parentheses.

TABLE 5

Learning-by-doing and Heterogeneity in Firm Performance: Robustness Checks^a

| Dependent Variable | Profit Dispersion | Profit Dispersion | Profit Dispersion | Profit Dispersion | q dispersion | q dispersion | q dispersion | q dispersion |
|--|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|---------------------|--------------------|
| Method of estimating learning coefficient | OLS | OP | ACF | FE-SURV | OLS | OP | ACF | FE-SURV |
| 1. 75 th to 25 th percentile as dependent variable | 0.717*** (0.10) | 0.426*** (0.08) | 0.430*** (0.06) | 0.213*** (0.08) | 1.34*** (0.13) | 0.789*** (0.11) | 0.593*** (0.08) | 0.130 (0.12) |
| 2. Standard deviation as dependent variable | 0.170*** (0.02) | 0.109*** (0.02) | 0.102*** (0.01) | 0.038** (0.02) | 1.32*** (0.14) | 0.902*** (0.13) | 0.738*** (0.09) | 0.151 (0.14) |
| 3. Learning ranks instead of coefficients ^b | 0.396*** (0.06) | 0.366*** (0.06) | 0.384*** (0.06) | 0.142 (0.09) | 0.649*** (0.08) | 0.582*** (0.07) | 0.510*** (0.06) | -0.101 (0.01) |
| 4. "High" learning vs "low" learning classification | 0.123*** (0.02) | 0.099*** (0.02) | 0.221*** (0.02) | 0.021 (0.02) | 0.187*** (0.03) | 0.139*** (0.03) | 0.1415*** (0.03) | -0.038 (0.03) |
| 5. Including only very focused firms ^c | 1.01*** (0.20) | 0.551*** (0.16) | 0.710*** (0.12) | -0.062 (0.15) | 1.05*** (0.24) | 0.512*** (0.18) | 0.853*** (0.14) | -0.148 (0.19) |
| 6. SIC-4 as industry definition (OLS only) | 0.452*** (0.11) | | | | 0.951*** (0.16) | | | |
| 7. Excluding inds. ending with "9" and including # of SIC-4 codes as control | 0.826*** (0.11) | 0.372*** (0.09) | 0.440*** (0.07) | 0.239*** (0.08) | 1.25*** (0.15) | 0.661*** (0.13) | 0.583*** (0.09) | 0.008 (0.13) |
| 8. Industry level instead of industry-year level ^c | 0.660* (0.35) | 0.353 (0.22) | 0.561** (0.22) | 0.044 (0.27) | 0.696 (0.49) | 0.273 (0.31) | 0.566* (0.31) | -0.552 (0.46) |
| 9. Without taking logarithms | 0.264*** (0.04) | 0.124*** (0.03) | 0.126*** (0.02) | -0.038 (0.03) | 2.00*** (0.32) | 1.11*** (0.25) | 0.916*** (0.17) | -0.642** (0.25) |
| 10. Dependent variable not logged | 0.351*** (0.04) | 0.197*** (0.03) | 0.188*** (0.02) | 0.072** (0.03) | 2.87*** (0.31) | 1.89*** (0.25) | 1.49*** (0.17) | 0.294 (0.27) |
| 11. Exclude industry-years with < 25 firms | 0.888*** (0.13) | 0.585*** (0.10) | 0.441*** (0.09) | 0.182** (0.09) | 1.26*** (0.16) | 0.972*** (0.13) | 0.659*** (0.11) | 0.233* (0.13) |
| 12. Cluster standard errors at SIC-3 level ^e | 0.926*** (0.21) | 0.485** (0.20) | 0.491*** (0.15) | 0.203 (0.26) | 1.41*** (0.27) | 0.803*** (0.27) | 0.677*** (0.17) | 0.027 (0.35) |
| 13. Exclude two highest and two lowest learning inds. | 0.726*** (0.16) | 0.523*** (0.12) | 0.399*** (0.08) | 0.175 (0.15) | 1.28*** (0.21) | 0.926*** (0.16) | 0.628*** (0.10) | -0.065 (0.21) |
| 14. Include SIC-2-year fixed effects | 0.974*** (0.14) | 0.380*** (0.12) | 0.408*** (0.08) | 0.010 (0.11) | 1.39*** (0.20) | 0.509*** (0.17) | 0.540*** (0.11) | -0.033 (0.16) |

| Dependent Variable | Profit Dispersion | Profit Dispersion | Profit Dispersion | Profit Dispersion | <i>q</i> dispersion | <i>q</i> dispersion | <i>q</i> dispersion | <i>q</i> dispersion |
|--|--------------------|-------------------|-------------------|-------------------|---------------------|---------------------|---------------------|---------------------|
| Method of estimating learning coefficient | OLS | OP | ACF | FE-SURV | OLS | OP | ACF | FE-SURV |
| 15. Time period 1973-1984 ^f (OLS only) | 0.640*** (0.17) | | | | 2.40*** (0.23) | | | |
| 16. Time period 1985-2000 ^f (OLS only) | 0.613*** (0.15) | | | | 1.50*** (0.21) | | | |
| 17. Learning coefficients with R&D controls (OLS only) ^g | 0.349*** (0.06) | | | | 0.633*** (0.08) | | | |

- a. This table provides the results of robustness checks that use the same type of regressions as in Table 4, but with different measures of performance heterogeneity, level of aggregation, choice of time periods, assumptions about error correlation structures etc. Each line in this table is comparable to line 1 from Table 4. Only the coefficients and standard errors on industry learning intensity are presented. Coefficients on other variables are available on request. "Method of estimating learning coefficient" refers to the estimation approach using to estimate the learning rates in Equation 3. OLS uses the OLS method. OP uses an extension of Olley-Pakes (1996), ACF uses an extension of Akerberg et al (2005) and FE-SURV uses plant fixed effects specifications on a sample of survivors.
- b. The coefficients and standard errors have been multiplied by 100 for presentation purposes.
- c. This regression includes only firms whose largest segment (by sales) accounts for at least 95% of the total sales. Due to non-availability of data before 1984, this covers the period 1984 to 2000.
- d. The unit of analysis in this regression is industry – hence, there are only 72 observations for the OLS, OP and ACF methods and 49 for the FE approach.
- e. The standard errors are computed allowing for arbitrary autocorrelation of errors within an SIC-3 industry.
- f. These use learning coefficients estimated for the relevant time period as an independent variable.
- g. Learning coefficients used in this regression are estimated after controlling for *firm specific* R&D expenditure (the sample changes because not all firms report R&D).

APPENDIX A: INDUSTRY-BY-INDUSTRY LEARNING COEFFICIENTS

| SIC | Rank | OLS | | OP | | ACF | | FESURV | |
|-----|-------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|
| | (OLS) | Coeff. | Std Error | Coeff. | Std Error | Coeff. | Std Error | Coeff. | Std Error |
| 201 | 99 | 0.148 | (0.02) | 0.228 | (0.05) | 0.253 | (0.11) | 0.208 | (0.06) |
| 202 | 79 | 0.180 | (0.02) | 0.268 | (0.08) | 0.222 | (0.09) | -0.018 | (0.08) |
| 203 | 18 | 0.321 | (0.02) | 0.436 | (0.06) | 0.503 | (0.05) | 0.122 | (0.05) |
| 204 | 30 | 0.273 | (0.02) | 0.326 | (0.06) | 0.546 | (0.06) | 0.109 | (0.06) |
| 205 | 23 | 0.294 | (0.02) | 0.491 | (0.08) | 0.243 | (0.05) | | |
| 206 | 58 | 0.207 | (0.04) | 0.332 | (0.13) | 0.396 | (0.12) | | |
| 207 | 17 | 0.323 | (0.04) | 0.327 | (0.10) | 0.289 | (0.11) | | |
| 208 | 21 | 0.300 | (0.02) | 0.389 | (0.06) | 0.488 | (0.08) | 0.204 | (0.08) |
| 209 | 38 | 0.260 | (0.02) | 0.395 | (0.04) | 0.296 | (0.05) | 0.080 | (0.08) |
| 221 | 101 | 0.145 | (0.07) | 0.372 | (0.14) | 0.305 | (0.17) | | |
| 222 | 77 | 0.181 | (0.04) | 0.140 | (0.14) | 0.339 | (0.21) | | |
| 224 | 98 | 0.151 | (0.08) | 0.056 | (0.18) | 0.156 | (0.22) | | |
| 225 | 72 | 0.187 | (0.02) | 0.245 | (0.04) | 0.300 | (0.06) | 0.212 | (0.07) |
| 226 | 94 | 0.163 | (0.03) | 0.221 | (0.07) | 0.247 | (0.10) | | |
| 227 | 16 | 0.326 | (0.05) | 0.348 | (0.18) | 0.262 | (0.19) | | |
| 228 | 116 | 0.041 | (0.03) | 0.098 | (0.07) | 0.158 | (0.12) | | |
| 229 | 55 | 0.213 | (0.03) | 0.421 | (0.08) | 0.178 | (0.07) | | |
| 231 | 26 | 0.284 | (0.04) | 0.341 | (0.15) | 0.371 | (0.14) | | |
| 232 | 92 | 0.165 | (0.02) | 0.238 | (0.04) | 0.182 | (0.05) | 0.212 | (0.09) |
| 233 | 29 | 0.278 | (0.01) | 0.415 | (0.04) | 0.401 | (0.02) | 0.116 | (0.08) |
| 234 | 46 | 0.241 | (0.04) | 0.298 | (0.08) | 0.152 | (0.07) | | |
| 235 | 32 | 0.271 | (0.08) | 0.202 | | 0.488 | | | |
| 236 | 40 | 0.254 | (0.03) | 0.235 | (0.12) | 0.306 | (0.13) | | |
| 238 | 35 | 0.266 | (0.04) | 0.456 | (0.10) | 0.220 | (0.10) | | |
| 239 | 25 | 0.286 | (0.02) | 0.394 | (0.04) | 0.304 | (0.06) | 0.215 | (0.09) |
| 241 | 56 | 0.213 | (0.01) | 0.352 | (0.04) | 0.220 | (0.05) | | |
| 242 | 64 | 0.202 | (0.01) | 0.288 | (0.03) | 0.275 | (0.03) | 0.003 | (0.06) |
| 243 | 61 | 0.205 | (0.01) | 0.263 | (0.03) | 0.223 | (0.06) | 0.150 | (0.06) |
| 244 | 110 | 0.093 | (0.02) | 0.198 | (0.05) | 0.244 | (0.06) | | |
| 245 | 109 | 0.108 | (0.02) | 0.198 | (0.03) | -0.005 | (0.13) | 0.049 | (0.05) |
| 249 | 82 | 0.177 | (0.02) | 0.253 | (0.04) | 0.265 | (0.05) | 0.188 | (0.06) |
| 251 | 87 | 0.170 | (0.01) | 0.306 | (0.03) | 0.130 | (0.06) | 0.024 | (0.06) |
| 252 | 67 | 0.198 | (0.03) | 0.370 | (0.07) | 0.292 | (0.09) | | |
| 253 | 76 | 0.182 | (0.04) | 0.311 | (0.08) | 0.316 | (0.09) | | |
| 254 | 96 | 0.154 | (0.02) | 0.308 | (0.07) | 0.224 | (0.06) | | |
| 259 | 53 | 0.215 | (0.03) | 0.378 | (0.06) | 0.230 | (0.11) | | |
| 262 | 115 | 0.054 | (0.04) | 0.239 | (0.20) | 0.166 | (0.25) | | |
| 265 | 69 | 0.192 | (0.01) | 0.336 | (0.03) | 0.337 | (0.07) | 0.160 | (0.04) |
| 267 | 37 | 0.260 | (0.03) | 0.236 | (0.14) | 0.486 | (0.08) | | |
| 281 | 20 | 0.310 | (0.02) | 0.427 | (0.08) | 0.531 | (0.08) | 0.054 | (0.07) |
| 282 | 9 | 0.385 | (0.03) | 0.466 | (0.09) | 0.465 | (0.07) | | |
| 283 | 2 | 0.547 | (0.03) | 0.800 | (0.09) | 0.718 | (0.09) | | |
| 284 | 10 | 0.372 | (0.03) | 0.430 | (0.12) | 0.542 | (0.10) | | |
| 285 | 15 | 0.342 | (0.03) | 0.574 | (0.07) | 0.589 | (0.20) | | |
| 286 | 8 | 0.388 | (0.03) | 0.565 | (0.14) | 0.524 | (0.11) | | |
| 287 | 6 | 0.413 | (0.03) | 0.505 | (0.12) | 0.531 | (0.10) | | |
| 289 | 5 | 0.413 | (0.02) | 0.485 | (0.04) | 0.536 | (0.04) | 0.154 | (0.06) |

| SIC | Rank | OLS | | OP | | ACF | | FESURV | |
|-----|------|--------------|------------------|------------------|------------------|------------------|------------------|--------|--|
| | | (OLS) | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error | | |
| 291 | 3 | 0.451 (0.10) | | 0.421 (0.34) | | 0.708 (0.34) | | | |
| 295 | 52 | 0.224 (0.02) | | 0.325 (0.05) | | 0.362 (0.05) | 0.064 (0.06) | | |
| 299 | 11 | 0.369 (0.05) | | 0.490 (0.2) | | 0.604 (0.31) | | | |
| 305 | 19 | 0.313 (0.05) | | 0.466 (0.14) | | 0.363 (0.11) | | | |
| 306 | 88 | 0.169 (0.02) | | 0.298 (0.07) | | 0.135 (0.08) | | | |
| 308 | 48 | 0.237 (0.01) | | 0.357 (0.02) | | 0.386 (0.03) | 0.147 (0.05) | | |
| 311 | 117 | 0.005 (0.07) | | 0.041 (0.23) | | 0.183 (0.18) | | | |
| 313 | 39 | 0.256 (0.09) | | | | | | | |
| 314 | 43 | 0.246 (0.03) | | 0.412 (0.09) | | 0.213 (0.10) | | | |
| 316 | 103 | 0.144 (0.08) | | | | 0.125 (0.17) | | | |
| 317 | 113 | 0.076 (0.07) | | 0.335 (0.20) | | -0.021 (0.16) | | | |
| 319 | 104 | 0.127 (0.07) | | 0.300 (0.30) | | -0.105 (0.26) | | | |
| 321 | 78 | 0.181 (0.05) | | 0.047 | | 0.370 | | | |
| 322 | 114 | 0.067 (0.03) | | 0.213 (0.10) | | 0.293 (0.13) | | | |
| 323 | 73 | 0.185 (0.02) | | 0.326 (0.07) | | 0.265 (0.09) | | | |
| 324 | 7 | 0.389 (0.08) | | 0.503 | | -0.100 (0.33) | | | |
| 325 | 12 | 0.360 (0.03) | | 0.524 (0.11) | | 0.445 (0.09) | 0.019 (0.12) | | |
| 326 | 97 | 0.152 (0.03) | | 0.234 (0.08) | | 0.238 (0.08) | | | |
| 327 | 36 | 0.265 (0.01) | | 0.358 (0.02) | | 0.415 (0.03) | -0.025 (0.08) | | |
| 328 | 31 | 0.272 (0.04) | | 0.294 (0.09) | | 0.126 (0.12) | | | |
| 329 | 14 | 0.354 (0.02) | | 0.501 (0.05) | | 0.323 (0.08) | | | |
| 331 | 90 | 0.169 (0.02) | | 0.290 (0.05) | | 0.280 (0.05) | 0.342 (0.07) | | |
| 332 | 85 | 0.172 (0.02) | | 0.267 (0.05) | | 0.279 (0.08) | 0.215 (0.10) | | |
| 333 | 45 | 0.241 (0.05) | | -0.091 | | 0.086 (0.17) | -0.160 (0.12) | | |
| 334 | 54 | 0.214 (0.05) | | 0.170 (0.11) | | 0.340 (0.09) | | | |
| 335 | 68 | 0.194 (0.02) | | 0.239 (0.04) | | 0.224 (0.04) | 0.297 (0.07) | | |
| 336 | 33 | 0.269 (0.03) | | 0.353 (0.06) | | 0.317 (0.08) | 0.005 (0.11) | | |
| 339 | 106 | 0.123 (0.03) | | 0.222 (0.09) | | 0.265 (0.10) | -0.035 (0.22) | | |
| 341 | 22 | 0.299 (0.02) | | 0.510 (0.08) | | 0.560 (0.07) | 0.099 (0.04) | | |
| 342 | 75 | 0.184 (0.02) | | 0.228 (0.04) | | 0.258 (0.06) | -0.090 (0.05) | | |
| 343 | 59 | 0.207 (0.03) | | 0.382 (0.12) | | 0.319 (0.10) | | | |
| 344 | 80 | 0.178 (0.01) | | 0.302 (0.03) | | 0.298 (0.03) | -0.001 (0.04) | | |
| 345 | 74 | 0.184 (0.02) | | 0.271 (0.06) | | 0.325 (0.10) | | | |
| 346 | 84 | 0.173 (0.02) | | 0.283 (0.03) | | 0.199 (0.04) | 0.061 (0.05) | | |
| 347 | 63 | 0.202 (0.01) | | 0.318 (0.07) | | 0.325 (0.05) | | | |
| 348 | 83 | 0.174 (0.05) | | 0.210 (0.17) | | 0.034 (0.22) | | | |
| 349 | 71 | 0.189 (0.01) | | 0.289 (0.03) | | 0.208 (0.04) | 0.095 (0.04) | | |
| 351 | 86 | 0.170 (0.04) | | 0.190 (0.12) | | 0.117 (0.11) | | | |
| 353 | 89 | 0.169 (0.02) | | 0.331 (0.03) | | 0.238 (0.06) | 0.038 (0.06) | | |
| 354 | 62 | 0.202 (0.01) | | 0.275 (0.02) | | 0.247 (0.04) | 0.150 (0.04) | | |
| 355 | 51 | 0.227 (0.02) | | 0.363 (0.05) | | 0.289 (0.07) | 0.012 (0.07) | | |
| 356 | 47 | 0.237 (0.01) | | 0.341 (0.03) | | 0.251 (0.03) | 0.214 (0.03) | | |
| 357 | 1 | 0.602 (0.03) | | 0.816 (0.06) | | 0.982 (0.05) | 0.658 (0.1) | | |
| 358 | 44 | 0.246 (0.02) | | 0.330 (0.05) | | 0.343 (0.05) | 0.071 (0.05) | | |
| 359 | 93 | 0.163 (0.01) | | 0.278 (0.02) | | 0.221 (0.04) | 0.053 (0.07) | | |
| 361 | 24 | 0.290 (0.02) | | 0.375 (0.06) | | 0.378 (0.07) | | | |
| 362 | 66 | 0.200 (0.02) | | 0.340 (0.09) | | 0.197 (0.08) | -0.019 (0.07) | | |
| 363 | 105 | 0.126 (0.04) | | 0.329 (0.09) | | 0.402 (0.13) | | | |
| 364 | 49 | 0.233 (0.02) | | 0.321 (0.06) | | 0.284 (0.09) | -0.028 (0.09) | | |
| 365 | 107 | 0.122 (0.04) | | 0.306 (0.10) | | 0.405 (0.11) | | | |

| SIC | Rank | OLS | | OP | | ACF | | FESURV | |
|---|------|-------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| | | (OLS) | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error | Coeff. Std Error |
| 366 | 50 | | 0.229 (0.03) | | 0.430 (0.08) | | 0.409 (0.07) | | 0.087 (0.10) |
| 367 | 13 | | 0.355 (0.02) | | 0.720 (0.06) | | 0.693 (0.07) | | 0.548 (0.04) |
| 369 | 28 | | 0.278 (0.02) | | 0.409 (0.11) | | 0.479 (0.06) | | 0.172 (0.07) |
| 371 | 70 | | 0.191 (0.01) | | 0.323 (0.03) | | 0.332 (0.04) | | 0.134 (0.03) |
| 372 | 95 | | 0.157 (0.02) | | 0.298 (0.05) | | 0.194 (0.09) | | 0.165 (0.06) |
| 373 | 108 | | 0.108 (0.02) | | 0.167 (0.04) | | 0.104 (0.04) | | |
| 374 | 102 | | 0.144 (0.05) | | 0.335 (0.14) | | 0.172 (0.10) | | |
| 375 | 91 | | 0.166 (0.06) | | 0.526 (0.21) | | 0.333 (0.15) | | |
| 379 | 112 | | 0.091 (0.02) | | 0.123 (0.07) | | 0.147 (0.09) | | |
| 381 | 27 | | 0.280 (0.07) | | 0.388 (0.12) | | 0.238 (0.09) | | |
| 382 | 81 | | 0.178 (0.02) | | 0.279 (0.03) | | 0.147 (0.04) | 0.236 (0.05) | |
| 384 | 34 | | 0.267 (0.02) | | 0.390 (0.04) | | 0.363 (0.07) | 0.139 (0.05) | |
| 385 | 111 | | 0.093 (0.04) | | -0.050 (0.14) | | 0.236 (0.12) | | |
| 386 | 4 | | 0.416 (0.04) | | 0.551 (0.11) | | 0.559 (0.15) | | |
| 391 | 42 | | 0.250 (0.03) | | 0.362 (0.06) | | 0.369 (0.07) | | |
| 393 | 100 | | 0.146 (0.04) | | 0.079 (0.11) | | -0.040 (0.16) | | |
| 394 | 41 | | 0.251 (0.02) | | 0.378 (0.08) | | 0.373 (0.06) | | |
| 395 | 65 | | 0.202 (0.05) | | 0.494 (0.28) | | 0.696 (0.15) | | |
| 396 | 57 | | 0.209 (0.04) | | 0.333 (0.14) | | 0.156 (0.15) | | |
| 399 | 60 | | 0.206 (0.02) | | 0.292 (0.04) | | 0.267 (0.06) | | |
| Mean | | | 0.227 | | 0.331 | | 0.308 | | 0.121 |
| Std. Dev | | | 0.097 | | 0.142 | | 0.172 | | 0.142 |
| Min | | | 0.005 | | -0.091 | | -0.105 | | -0.160 |
| Max | | | 0.602 | | 0.816 | | 0.982 | | 0.658 |
| Spearman' Rank Correlation Coefficient (<i>p-values in parentheses</i>) | | | | | | | | | |
| OLS | | | 1 | | 0.74(0.00) | | 0.63(0.00) | | 0.09(0.54) |
| OP | | | | | 1 | | 0.57(0.00) | | 0.12(0.41) |
| ACF | | | | | | | 1 | | 0.18(0.21) |

Notes:

- (1) Rank (OLS) represents the rank based on the OLS learning coefficient.
- (2) "OLS" uses the OLS method. "OP" uses an extension of Olley-Pakes (1996), "ACF" uses an extension of Akerberg et al (2005) and "FE-SURV" uses plant fixed effects specifications on a sample of survivors that survived at least till age 8.
- (3) Standard errors presented in parentheses. Standard errors for OP and ACF methods were generated by bootstrapping the sample 50 times (the resampling unit was a plant, not a plant-year).
- (4) Missing cells indicate that there were less than 50 plants or that the sample size was too small to meet Census disclosure requirements.

APPENDIX B: ROBUSTNESS CHECKS

Are we measuring learning-by-doing?

We tested if there are other phenomena that manifest as a high coefficient on cumulative output. While we cannot rule all possible alternatives out, we attempt to identify and argue against some important ones.

Survivor Bias and Sample Selection: With OLS, endogeneity of exit may bias the measured coefficient on cumulative output. If accumulated experience helps firms withstand bad “performance shocks” and thus reduces the probability of exit, then the measured coefficient will be biased downward. On the other hand, if accumulated experience has no effect on exit, then the bias may be upward, the argument being that only “good” firms survive and they would tend to have both higher cumulative output and higher productivity, resulting in a high “learning” coefficient. However, these arguments does not appear to hold in our study for several reasons. First, the Olley-Pakes and ACF approaches make explicit assumptions about selection, and within the bounds of these assumptions the estimates correct for selection bias. Next, if the second argument were true, we should see a strong positive correlation between the turnover rate of firms (entry rate + exit rate) and our measured “learning”. However, we see no statistically significant relationship between industry turnover rate and the measured learning rate (results available on request). Third, using learning estimates from a fixed effects specification on a sample of survivors seems to support some of our subsequent results on the heterogeneity of firm performance.

We also tested the robustness of our results to the choice of our sample. Using a sample that includes only ASM plants (which have better quality data) and relaxing our condition that plants not have a gap of more than 2 years between two consecutive years produced learning estimates that were highly correlated with our baseline estimates (results available on request).

R&D investments: Sinclair et al (2000) argue that it is specific R&D efforts that cause

learning-by-doing. If all the learning were due to R&D, we should observe no learning once we include R&D as a control. Without controlling for R&D, we would observe high R&D industries to have high learning. Since the industries with high measured learning in our study are R&D intensive industries, we attempt to rule out R&D as *solely* driving the results. However, we lack detailed data on plant-level R&D and hence, we rely on a more rudimentary check. We use firm-level R&D from Compustat as controls and re-estimate the learning coefficients. The rank correlation between this set of learning coefficients and our original estimates is 0.67 and statistically significant at the 0.01% level.¹⁷ Furthermore, using these revised learning coefficients did not change the subsequent results on firm performance heterogeneity (Row 17 in Table 5).

Measurement errors in capital: It is well known that there are errors with measuring capital. If it were true that such measurement errors were more prevalent in some industries (e.g. in hi-tech industries), then we may observe a high measured rate of learning in such industries. While there is no known way to completely rule this out, we re-estimate our learning coefficients using an alternative measure of capital instead of the perpetual inventory method used in the study. Specifically, we use the year-end book value of assets and find that the resulting learning coefficients are highly correlated with our original estimates.

Industry life cycles: The need for learning-by-doing may be intricately linked to industry life cycles. For instance, early in the industry life cycle, firms may need to learn mostly on their own. Furthermore, this is also a period of great uncertainty and consequently higher heterogeneity of performance. As the industry matures, dominant design(s) emerge and firms may be able to benefit from others, thereby reducing the need for own learning. Concomitantly, the uncertainty also decreases, reducing performance heterogeneity. While this is certainly consistent with our arguments (note that we do not make any exhaustive claims over what gives

¹⁷ Please note that the sub-sample with R&D data is much smaller than the overall sample.

rise to learning-by-doing), that we find exit rates to be uncorrelated with learning intensities suggests that industry life cycles are not the sole driving factor here.

Alternative production functions: We tested the robustness to relaxing the Cobb-Douglas production function form adopted here. Specifically, we estimated using OLS, a version of the Translog production function ($y_{ijt} = \alpha_{jt} + \alpha_j \cdot k_{ijt} + \alpha'_j \cdot (k_{ijt})^2 + \beta_j \cdot l_{ijt} + \beta'_j \cdot (l_{ijt})^2 + \gamma_j (k_{ijt})(l_{ijt}) + \lambda_j \cdot x_{ijt} + \varepsilon_{ijt}$). These learning coefficients were highly correlated with our baseline estimates.

Learning-by-doing and heterogeneity of firm performance

Table 5 presents the results of robustness checks on this aspect.

Alternative measures of heterogeneity: So far, we have used the difference between the 90th and 10th percentiles as the measure of firm performance heterogeneity. Rows 1 and 2 use the inter-quartile range (difference between the 75th and 25th percentiles) and standard deviation of firm performance respectively. We observe that the coefficient on learning intensity is strongly positive in all but one specification.

Ordinal measures of learning intensity: In our baseline results, we directly used the estimated learning intensities as measures of learning-by-doing. Since the estimated learning intensities vary across different specifications, we check if using ordinal measures makes a difference. Row 3 uses the industry learning ranks (based on the estimated learning coefficient) instead of the estimated learning intensities. Row 4 uses a simple dummy variable that divides industries into “high learning” and “low learning” based on the median estimated learning coefficients. Once again, the results are statistically significant in almost all specifications.

Excluding diversified firms: Our results use the firm’s primary SIC code to assign firms to industries. However, many firms in Compustat are diversified and hence, it may be argued that it is inappropriate to use a single learning coefficient for such firms. Using the Compustat business segments data from 1984 to 2000, we select firms that have a single SIC-3 segment that comprises at least 95% of their total sales and estimate Equation 5 for these firms. Row 5 presents the results. The results using learning estimates from all non-fixed effects

specifications are strongly positive but the regressions using learning coefficients from fixed-effects models become statistically insignificant. We must, however, note that when using this approach, the number of industry-year cells with adequate number of firms falls considerably.¹⁸

Level of aggregation: Since it may be argued that SIC-3 digit is a very high level of aggregation, we estimate learning intensity at the SIC-4 digit level using OLS (the sample size is too small to let us use the other specifications) and repeat our test.¹⁹ Row 6 shows that even with a finer industry definition, our results hold. Another possible concern is that the definition of some SIC-3 industries are simply much more heterogeneous than others. To rule this possibility out, (a) we excluded all SIC-3 industries ending with “9” (usually “not elsewhere classified” industries) and (b) included the number of SIC-4 digit codes within a SIC-3 as a control in Equation 5. The results remained statistically significant (Row 7) for most specifications.

Other econometric concerns: Rows 8 to 17 provide results to alleviate other possible econometric concerns such as the use of industry-year as a unit of analysis (check: industry as a unit of analysis), errors being correlated within an industry (check: cluster errors within an industry), taking logarithms (checks: not taking logarithms at all and not taking logarithms for the dependent variable), the choice of 10 firms in an industry-year as the cut-off for inclusion (check: use 25 firms as cut-off), influence of outliers (check: exclude the top two and bottom

¹⁸ A related concern might be that not all plants used to estimate industry learning rates belong to Compustat firms and hence, that it might be problematic to apply the learning rates from Census data to Compustat. We estimated learning coefficients using only a (much smaller) sample of plants that could be matched to Compustat firms and found that even based on these coefficients, the relationship between learning-by-doing and heterogeneity of firm performance remained positive and statistically significant.

¹⁹ The rank correlation between the SIC-3 estimates of learning and the weighted average SIC-4 estimates belonging to the same SIC-3 industry is about 0.88.

industries in terms of learning rates, choice of time period (check: separate 1973-1984 and 1985-2000), other unobserved industry factors (check: *joint* SIC2-year dummies). The results remain statistically significant for a major part, though not all, of these alternative specifications. Please note that sample size constraints prevented us from using non-OLS approaches in some cases. To summarize, our results appear to hold under a variety of different specifications and it seems reasonable to conclude that we are measuring learning-by-doing, at least in a rank-order sense.

APPENDIX C: NOTE ON OP AND ACF ESTIMATION APPROACHES

The basic OP and ACF methods (without any experience term) write the production function as:

$$v_{ijt} = a_{jt} + \alpha_j \cdot k_{ijt} + \beta_j \cdot l_{ijt} + \omega_{ijt} + \theta_{ijt}$$

where ω_{ijt} is a *and the only* heterogeneity that is not known to us but known to the firm and incorporated into its decisions and θ_{ijt} is a completely exogenous error term. Both approaches assume that ω_{ijt} can be written as:

$$\omega_{ijt} = E[\omega_{ijt} | \omega_{ijt-1}] + \xi_{ijt} \text{ where } E[\xi_{ijt}] = 0 \quad (C1)$$

An AR(1) process would be a special case of C1. Suppose, we could somehow estimate ω_{ijt} , and hence, ξ_{ijt} . Then, we could make assumptions about when firms decide on their inputs to get identifying moment conditions. Specifically, these approaches assume that capital at period t is completely decided at period $t-1$ when the firm knows only ω_{ijt-1} and the *expected* value of ω_{ijt} , which then implies that the capital decision at time t is uncorrelated with the *unexpected* change ξ_{ijt} . So, we have

$$E[\xi_{ijt} | k_{ijt}] = 0 \quad (C2)$$

This can then be used to identify the capital coefficient α_j using GMM estimation.

In order to estimate the learning coefficient, we add a cumulative output term (x_{ijt}) to the above equation and assume that the prior cumulative experience of a firm at time t is completely determined at time $t-1$. Based on the reasoning above, it follows that

$$E[\xi_{ijt} | x_{ijt}] = 0 \quad (C3)$$

which can be used to identify the learning coefficient λ_j .

To estimate ω_{ijt} , the OP method assumes that an observable decision (e.g. capital investment) is a strictly increasing function of ω_{ijt} and capital i.e. $i_{ijt} = f(\omega_{ijt}, k_{ijt})$, which can be inverted to “back out” the unobserved heterogeneity i.e. $\omega_{ijt} = f^{-1}(i_{ijt}, k_{ijt})$, f^{-1} approximated using a higher-order polynomial. In our case, we assumed that firm capital investment will also be a function of cumulative experience and hence, that $\omega_{ijt} = f^{-1}(i_{ijt}, k_{ijt}, x_{ijt})$. The ACF approach adds labor in addition to the other terms into this function. We largely followed the original papers

with regard to implementation. The specific steps used to estimate the learning coefficient by the two approaches are given below. To avoid clutter, we drop the subscripts 'i' and 'j'.

OP Method

Step 1: Obtaining the labor coefficient: Regress using OLS, y_t on labor (l_t), and a third degree polynomial of i_t , k_t and x_t . This gives the estimated labor coefficient β^* .

Step 2: "Backing out" the unobserved heterogeneity: Obtain the predicted values of $(y_t - \beta^* \cdot l_t)$, which are unbiased estimates of $\alpha k_t + \lambda x_t + \omega_t$. Call these predicted values ϕ_t^* .

Step 3: Selection correction: Obtain the predict values from a probit regression of an exit dummy on a third degree polynomial of i_t , k_t and x_t . Call these predicted values P_t^* .

Step 4: Estimate the capital and learning coefficients: Use a non-linear regression of the following form: $\phi_t^* = \alpha k_t + \lambda x_t + \text{third-degree polynomial in } (\phi_{t-1}^* - \alpha k_{t-1} - \lambda x_{t-1}) \text{ and } P_{t-1}^*$

ACF Method

We use the decision on materials expenditure m_t to invert out the unobserved heterogeneity. We assume labor inputs l_t are decided at $t-1$ and that the exit decision is made at $t-1$ (which implies we do not need a separate selection correction step as above)

Step 1: Regress using OLS, y_t on a polynomial of l_t , m_t , k_t and x_t . Obtain predicted values of y_t , which are unbiased estimates of $\alpha k_t + \beta l_t + \lambda x_t + \omega_t$ (So, no coefficients are estimated here)

Step 2: Choose some initial set of parameter values for α , β and λ . Using the predicted values from Step 1 and the chosen parameter values, obtain the estimated ω_t .

Step 3: Regress using OLS, the ω_t from Step 2 on a polynomial in ω_{t-1} . The residuals from the regression give the estimated ξ_t in Equation C1.

Step 4: Using the estimated ξ_t , say ξ_t^* , construct the empirical equivalent of the moment conditions C2 and C3 i.e. $(1/N)(\sum k_t^2 (\xi_t^*)^2 + x_t^2 (\xi_t^*)^2)$ where N is the sample size.

Step 5: Check if the sum of squares in Step 4 is close to zero. If yes, the parameter values in Step 2 are the estimated coefficients. If not, change parameter values in Step 2 and repeat till the condition is achieved.